1976

Assessing Inferential Accuracy In Clinical Judgment And Person Perception

Philip Lee Reed

INFORMATION TO USERS

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

This copy was produced from a microfiche copy of the original document. The quality of the copy is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

Canadian Theses Division
Cataloguing Branch
National Library of Canada
Ottawa, Canada K1A 0N4

AVIS AUX USAGERS

LA THESE A ETE MICROFILMEE TELLE QUE NOUS L'AVONS RECUE

Cette copie a été faite à partir d'une microfiche du document original. La qualité de la copie dépend grandement de la qualité de la thèse soumise pour le microfimage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

NOTA BENE: La qualité d'impression de certaines pages peut laisser à désirer. Microfilmée telle que nous l'avons reçue.

Division des thèses canadiennes
Direction du catalogage
Bibliothèque nationale du Canada
Ottawa, Canada K1A 0N4

ASSESSING INFERENTIAL ACCURACY IN CLINICAL JUDGMENT AND PERSON PERCEPTION

by

Philip Lee Reed

Department of Psychology

Submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Faculty of Graduate Studies
The University of Western Ontario
London, Ontario

June, 1976

Page

Of particular interest were the correlations between the two objective criteria (exemplars and modal profiles) and the group consensus judgment which ranged from .44 to .85. Partialing out the effects of desirability or base rate often raised these substantially. In addition, the lowest group consensus generalizability coefficient was .92. Study II found further evidence for the generalizability of group consensus accuracy.

Study III identified components of the group consensus judgment using a multiple correlation strategy. Content, desirability, and base rate accounted for an average of approximately 70% of the variance across 8 studies, with content being more important than the other two components combined.

Study IV successfully demonstrated the generalizability of individual sensitivity scores across targets and domains of behaviour. Distributions and statistics for all parameters are appended.

The results of these investigations are interpreted as providing support for certain basic ideas underlying this research and other related studies. These include the conceptualization of behaviour in terms of patterns of individual behavioural consistency, using a model of measurement concerned with dimensions of behaviour, and representing the patterns of consistency in terms of profiles of behaviour probabilities. In addition, the strength and consistency of the reported relations strongly indicate the usefulness of the concept of a multidimensional network of relations among behaviours. The function of this network in the process of making inferences or in predicting the behaviour of others has been presented as a central feature of the model for inferential accuracy. Accordingly, the appropriateness of this model is substantiated by these results, particularly with respect to the assessment of accuracy.

ACKNOWLEDGEMENT

A great deal is owed those who played a role in the completion of this dissertation. In particular, I thank Dr. Douglas N. Jackson, who in serving as Chief Advisor provided the highest level of intellectual stimulation and encouragement.

I also thank Dr. R. C. Gardner, Dr. R. W. J. Neufeld, Dr. D. S. Pepper, and Dr. I. Spence for serving as members of the Advisory Committee, and my colleagues, Dr. Harvey A. Skinner and Jack Miedzinski, for their helpful comments and advice.

Financial support was provided by Canada Council Grants Number S74-0761, Douglas N. Jackson, Principal Investigator, and Number S75-0495-X1, Douglas N. Jackson and Philip L. Reed, Co-Principal Investigators.

Special thanks go to Elizabeth Du Maresq for her careful typing of the manuscript, and to Bernadette Furman Benko and Mary Jane MacKay for their help in the preparation of tables and appendices.

In addition, I wish to thank Dr. Arthur DeLong and Dr. Bernard E. Lyman for their very significant contributions made long ago.

Finally, and most important, I thank Gail for seeing a different way, and staying until this was finished.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

Emerging traditions in the fields of personality, personality perception, and human judgment are notable for their emphasis of one or more of three major concepts. The first is that behaviour is essentially inconsistent because it is primarily determined by situational factors and therefore changes as the situation changes (Mischel, 1968, 1973). The second concept is that individuals have cognitive or conceptual dimensions of behaviour such as traits, and further that individuals force consistency on what are basically empirically inconsistent data by using these personal constructs cognitively to organize perceived behaviour (Chapman & Chapman, 1971; Jones & Nisbett, 1972; Mischel, 1968, 1973, Schneider, 1973). The third concept involves proposals of models of human judgment stressing that individuals typically use irrational, inaccurate, or considerably suboptimal strategies of judgment (Chapman & Chapman, 1971; Kahneman & Tversky, 1973; Nisbett & Borgida, 1975).

On the other hand, during this same period of time an alternative tradition has developed suggesting that the use of modern assessment techniques will overcome many of the difficulties which stimulated these approaches (Jackson, 1971, 1975). This alternative tradition proposes a model of behaviour in which behaviour consistency and rational judgment are basic assumptions.

This dissertation has several aims. The first is to explicate in detail the basic models for the measurement and perception of behaviour consistency which underly this alternative tradition. Second is the presentation of empirical and logical evidence evaluating the validity of the three concepts presented above. A third purpose is to present a model originally developed by Jackson (1972; Reed & Jackson, 1975) concerned with how individuals make accurate and rational inferences about the behaviour of others. A fourth major aspect of the dissertation is the presentation of several modifications of this model for inferential accuracy. Related to this is a comparison of the components of accuracy scores generated by this model with those proposed by Cronbach (1955) in what has for some time been considered a definitive analysis of social perception scores. Finally, because the usefulness of the model depends on the availability of valid criteria, a comprehensive empirical analysis of criteria for assessing inferential accuracy has been conducted, and is reported in Chapter Eight.

CHAPTER TWO

TRAIT INFERENCE

Particularly relevant to the present discussion is the term "trait inference." It was first suggested by Bruner and Taguiri (1954), and by Cronbach (1955), that individuals have naive expectations regarding networks of implicative trait relations. Jackson (1962, 1969) has suggested that these expectations are best conceptualized as an n-dimensional Euclidean space in which traits are organized in terms of their implicative relations, which in turn are based on the actual co-occurrence of traits. In other words, a multidimensional model is postulated to represent both the cognitive and empirical organization of trait co-occurrence. Jackson (1972) calls this organization of traits the trait inferential network.

One early research effort which is quite relevant to this discussion is the well-known work of Cattell (1946), who investigated the structure of adult peer ratings. This approach uses factor analysis to determine the dimensions underlying the intercorrelations of bipolar ratings of persons known to the raters. Reaching a more refined state with Tupes and Christal (1961), and Norman (1963), in general the work up to and including these studies was considered to reflect the actual structure of personality traits as they existed in the ratees and was not considered to explore the cognitive inferential network or implicit personality theory of the raters. However, this interpretation was soon challenged, and it was suggested rather that the dimensions of trait attribution or trait ratings reflected primarily culturally shared connotative meaning (D'Andrade, 1965; Mulaik, 1964), and the personal constructs of the observer (Mischel, 1968, 1973).

Arguments Concerning Biases in the Perception of Others

Schneider (1973) has aptly labeled this controversy the "realism issue." The question is whether ratings in inferential studies reflect veridical perceptions of the actual co-occurrence of personality traits in stimulus persons, or if rather these ratings only reflect shared semantic meaning, idiosyncratic experiences, or a combination of the two. A great deal of data has been collected bearing on this issue. For example, Mulaik (1964) reported that factors derived from semantic ratings of traits are very similar to factors derived from ratings of people on the same traits. D'Andrade (1965) showed that a factor analysis of ratings of similarity of the meaning of trait names produced a factor structure that was very much like that found by Norman (1963) based on ratings of people. Passini and Norman (1966) found that ratings of strangers yielded the same structure as ratings of acquaintances.[1]

The process proposed to be responsible for the apparently paradoxical findings of Passini and Norman is actually quite straightforward. Typically it is assumed that the perceivers' shared implicit theory of personality produces similar covariations in ratings. That is, once the rater makes an assumption about the stimulus person with respect to a trait, ratings on the other traits are logically

[1]It is relevant here to question the ecological validity of these instances when raters have been asked to make judgments about perfect strangers. While researchers should always watch for circumstances where biases in judgments may be present due to insufficient information, in many instances the importance and ubiquity of this phenomenon have been exaggerated.

determined. Another way of putting it is that knowledge of the network of implicative relationships among traits makes it possible for a rater to consistently infer behaviours, once limited assumptions have been made about the person being rated. Explanations invoking similar processes with respect to semantic meaning are offered for both the Mulaik and D'Andrade results. Other authors (Mischel, 1968, 1973, Schneider, 1973) have suggested that a process of this nature was responsible for all consistency in trait measurement, including both ratings of close acquaintances and self-report personality questionnaire results.

Norman and Goldberg (1966) were not satisfied that an explanation of the type given above was adequate, especially with respect to ratings among those who are well acquainted. To explore the matter further they used a Monte Carlo procedure to generate a set of hypothetical ratings which could be assumed to be determined only by a "shared network of implicative trait relations." The ratings had the properties of being unrelated to the hypothetical targets being rated (zero correlations between raters) but had the same internal structure (identical correlations between scales for each "rater"). Thus, the structure revealed by a factor analysis of these Monte Carlo data could be said to represent a shared set of implicative trait relations only.

Norman and Goldberg compared these data with other ratings from groups varying in the length of acquaintance of the group members, and found consistent external structure across groups. However, they found that inter-rater agreement was directly related to degree of acquaintance. In other words, the greater the amount of knowledge about the ratees, the higher the agreement among raters. This, in conjunction with their finding that for the well acquainted groups the factor scores had greater convergent and discriminant validity, suggests that at least to some degree peer ratings reflect veridical perception of stimulus person behaviour.

Lay and Jackson (1969) lend further support to this position with their finding of "a remarkably close correspondence" between the structures of inferential judgments and actual personality scale responses. Using personality items representative of traits, these authors determined through multidimensional scaling the dimensions of judgments of the probability of joint endorsement of items. The resulting dimensions were then compared to the dimensions of factor analyses of independent groups' actual responses to the PRF. The reported extreme similarity of the two structures was interpreted as implying that trait-inferential judgments reflect veridical perception of actual behaviour covariation. Stricker, Jacobs, and Kogan (1974) report a similar finding using MMPI items from the Pd scale, and a substantially different method of comparing the structures of the two sources of data. In addition, Lay, Burron, and Jackson (1973) published further evidence supporting this interpretation.

The conclusion that these data reflect veridical perception of behaviour covariation has been criticized by Schneider (1973), who argues that the responses to personality items are no less a measure of implicit personality theory than are the trait-inferential judgments. He states, "When subjects filled out the questionnaires, they acted as perceivers of their own behaviour; Hence the measures of 'real' personality may also be biased by implicit personality theory." (Schneider, 1973, p. 302.) Schneider is saying that the structure and consistency found in self-report personality inventories is the product of a distorting cognitive process. He suggests this process is similar to that operating in studies such as Mulaik (1964) and D'Andrade (1965), however, in this case it is the individual's perception of his/her own behaviour that is being distorted. As long as this possibility is not ruled out, the extreme similarity of structure for Lay and Jackson's two data sets cannot be assumed to result from veridical perception of behaviour. This is an extremely important matter, as the explanation of these data is central to the entire controversy. Therefore the arguments will be examined in some detail.

Arguments Concerning Biases in the Perception of One's Own Behaviour

The process through which self-report personality data are alleged to be biased by implicit personality theory has not been as clearly explicated as in the case of ratings of others' behaviour. Apparently the argument is that as an individual responds to each statement, he consistently adjusts his responses such that two criteria are satisfied. The first criterion is that the responses, or more precisely their behavioural referents, be consistent with the respondent's understanding of the trait inferential network; the second criterion is that any given response be consistent with those already given previously in the test. This latter criterion implies that the order of the items could have profound effects on the test results, and would constitute a type of within test secular trend mentioned by Loevinger (1957). There is simply no evidence with respect to the PRF (Jackson, 1967, 1974) suggesting the presence of secular trends. The distortions suggested which are related to the first criterion are also logically inconsistent with certain empirical evidence. Interestingly, it is evidence usually cited in support of a situationist position. This research was carried out within the context of investigating causal attribution.

In an influential book of readings on attribution theory (Jones, Kanouse, Kelly, Nisbett, Valins, & Weiner, 1972) Jones and Nisbett (1972) discuss several studies examining what they describe as "the contention that actors and observers differ fundamentally in the processing of available data." Specifically, they have found that actors tend to attribute behaviours to situational demands, while observers tend to attribute the same actions to stable underlying personal dispositions. In their view, actor - observer differences in information processing "...exist for the basic reason that different aspects of the available information are salient for actors and observers and this differential salience affects the course and outcome of the attribution process." (Jones and Nisbett, p. 85, italics in original). Their position is basically that actors attend to situations and that observers attend to behaviour.

This reported tendency for actors to view their own behaviours as situationally determined has profound implications for the argument that implicit personality theories are responsible for a bias in an individual's perceptions of his own behaviour. To repeat, it has been suggested that implicit personality theory biases perceptions toward being more consistent. The causal attribution research, however, indicates that individuals in fact tend to view their behaviour as situationally determined. This implies that any bias in self report measures would be caused by attending too much to situations rather than to the relationships among behaviours. Recall that the reason suggested by Jones and Nisbett for an actor's attributing the causes of behaviour to situations is the high saliency of the situation in the actor's perception of his own behaviour. Thus, it follows that it is clearly not true that the primary bias in the perceptions and self reports of one's own behaviour is toward consistency. In fact, if one were to follow the logic of a situationist view of behaviour, situational determinants would lead to inconsistency in behaviour, and therefore any bias would be in the direction of over attributing inconsistency, rather than consistency.

The above arguments have ramifications both for the specific issue at hand, the veridicality of perceived behavioural co-occurrence, and for the general area of personality assessment. With respect to personality assessment, these logical and empirical data indicate that one can be confident that reported consistencies in behaviour, and stability in the structure of the covariation of behaviours are not significantly influenced by perceptual biases. Regarding the veridicality issue, one can also feel comfortable in accepting studies such as Stricker et al. (1974), Lay and Jackson (1969), and Lay et al. (1973), as evidence supporting veridical perception of behavioural co-occurrence and covariation.

There are additional reasons why one should take seriously the findings and conclusions of the three studies mentioned directly above. These have to do primarily with evidence for the generalizability of peer ratings. In a summary of over twenty studies dealing with this, Wiggins (1973, p. 356) concluded that "when certain technical considerations are taken into account, peer ratings have substantial generalizability across both observers and situations observed." Wiggins points out that peer ratings have been found to be useful predictors of various types of criteria, including military officer effectiveness, performance in flight training, leadership, disciplinary problems, teacher effectiveness, Peace Corps volunteer performance, the performance of supervisors in industry, and academic performance. Each of the citations made by Wiggins can be viewed as an instance in which peers demonstrated relatively accurate perception of stimulus person attributes. Certainly these findings cannot be ignored in our determination of the veridicality of observers' perceptions of consistency in behaviour.

Additional Instances of Accuracy in Social Perception

An additional consideration which should not be ignored with respect to the veridicality issue, is the type of perceptual accuracies which are acknowledged by those objecting to the idea of accuracy in peer ratings (i.e., social learning theorists). In this regard it is interesting to note that social learning theory and its accompanying methodology are based on assumptions of various types of perceptual and inferential accuracy. One of these is the assumed accuracy of behavioural observation. Another is perceptual and inferential accuracy with regard to subtle situational determinants of behaviour. A third type of accuracy is assumed with regard to an individual's anticipation of the consequences of behaviour. This is an especially interesting type of accuracy because it involves an inferential judgment about the probable consequences to oneself based on an observation of another's behaviour, or one's own past behaviour in the case of direct reinforcement. As Bandura puts it, "In most instances customary outcomes are reasonably good predictors of behaviour because the consequences that people anticipate for their actions are accurately derived from, and therefore correspond closely to, prevailing conditions of reinforcement." (Bandura, 1971, p. 36).

Bandura's position is both sound and reasonable. He is maintaining that this type of perception and inferential judgment is generally, but not always accurate. It seems equally sound and reasonable to maintain in a parallel fashion that the perception of behavioural consistency is generally, but not always accurate. In the same way that a person must correctly perceive reinforcement contingencies in order to function efficiently, he/she must correctly perceive behavioural patterns (unless, of course, reinforcement is never contingent on such accurate perception). In other words, there is a functional reason for the accurate perception of behaviour, and we can expect people to use a consistency or trait model of man only if it is accurate and adaptive. Since the evidence indicates strongly that people nearly universally use this model in describing others, it seems difficult to avoid the conclusion that it is accurate and adaptive.

Summary

This chapter has reviewed evidence and presented arguments relevant to the questions of the veridicality of peer ratings, trait inferences, and self reports. It has been shown that empirical data do not support the presence of a bias toward consistency in self reports of behaviour, and that any bias which is present most likely is in the direction of inconsistency. Therefore, evidence of consistent structure in behavioural styles found in self report data should be taken quite seriously. Furthermore, the high level of congruence between the structures of personality data and trait inference data indicates that at the group level accurate perception of the empirical relationships among traits is the rule. This does not mean that raters always attribute characteristics accurately to those they are rating. Passini and Norman, and Norman and Goldberg have shown that there exists considerable variability with respect to accuracy, and that not surprisingly, validity is associated with degree of acquaintance of rater with ratee. This indicates that precautions should be taken to assure that confidence is not placed in peer ratings obtained in contrived rating situations, and instances where raters are inadequately acquainted with the ratees. In general, however, the empirical and logical evidence suggest that in the majority of realistic situations, there is considerable accuracy in peer ratings and perceptions of others.

One final comment should be made before moving on to a discussion of clinical judgment research. Citing the D'Andrade, Mulaik, and

Passini and Norman studies, the assertion has been made repeatedly by others that it is particularly damaging to a consistency position that there is a stable structure present in peer ratings even in the absence of relevant information about ratees (Mischel, 1968, 1973; Schneider, 1973). This interpretation is puzzling, for it would seem that one would expect logical and consistent relationships among the terms used to refer to personality characteristics and their behaviour referents. To have found unstable structure would have been the embarrassing outcome for a theory of behaviour consistency. It would certainly be fallacious logic to assert that a theory could be disconfirmed by empirical findings which are consistent with the theory. Nonetheless, this is precisely what has been suggested. A paradoxical argument has been offered suggesting empirical data reflecting high levels of behavioural and perceptual consistency disprove a theory based on assumptions of consistency. In response to this assertion, one can only reply that it seems clearly evident that research in the area of implicit personality theory and trait inference has produced results supporting the underlying psychological theory, and therefore alternative lines of research will have to be relied on by critics of theories of behavioural consistency. One such alternative might be the frequently cited body of research dealing with clinical judgment.

CHAPTER THREE

CLINICAL JUDGMENT

In the area of clinical judgment research one finds several issues of controversy, including actuarial vs. clinical prediction, linear vs. configural judgmental strategies, and the accuracy of clinical assessment. Excellent reviews of these issues have appeared elsewhere (Goldberg, 1968; Sawyer, 1966; Wiggins, 1973), and because the research findings have been unusually consistent, it is not necessary to review the studies in any detail. It is sufficient to repeat here that, in general, actuarial prediction is consistently superior or equal to clinical prediction, there exists only very limited evidence of configural strategies in judgment, and the accuracy or validity of the clinician has been somewhat less than impressive (Goldberg, 1968; Wiggins, 1973).

Meaningfulness in Clinical Judgment Tasks

While there can be little argument over the trend of research findings to date, the implications of the findings are less clear. In fact, it is the contention of this author that nearly all of the clinical judgment studies conducted thus far have failed to present the judges with a psychologically meaningful task. In this author's view, a task is meaningful if the judge is asked to assess or predict the co-occurrence of behaviours or characteristics for which empirical evidence is available indicating that such co-occurrence does in fact take place. In short, there can be no valid human judgments nor a useful model thereof, if such judgments concern empirical relationships which are nonexistent, extremely weak, or highly contrived. The position taken here implies that a task cannot be psychologically meaningful if there exists no possible basis for direct or indirect experience with the relationships in question. If it is indeed the case that most tasks have lacked meaningfulness, it is not surprising that judges have faired poorly in their attempts at clinical prediction and judgment.

The necessity of developing a meaningful judgment paradigm is central to the model for inferential accuracy which is presented below. Therefore, it will be helpful to consider factors that have contributed to the reduction of psychological meaningfulness in various clinical judgment studies, bearing in mind that level of meaningfulness is tied to the strength of underlying empirical relationships.

Working within the context of the Lens Model, Hursch, Hammond, and Hursch (1964) have developed algebraically several principles which should also apply to the present general analysis of clinical judgment. Hursch et al. demonstrated that mathematical limits on the maximum possible achievement by a judge are imposed by (1) the empirical relationships existing among the cues, (2) the empirical relationships between cues and criteria, and (3) whether or not the measurement model being used allows for nonlinear relationships. The relationships in question are generally expressed in terms of the correlations between cues and criteria, and the correlations among cues. These, in turn, are partially a function of the reliability with which both the cues and criteria can be assessed. Both cue and criterion reliability are mentioned, for while it is generally recognized that low cue or predictor reliability can attenuate correlations, the co-equal role played by criterion reliability is sometimes overlooked.

As asserted above, most clinical judgment studies provide excellent examples of one or more sources of meaning reduction. Regarding cue reliability, for example, many studies (i.e., Golden, 1964; Grebstein, 1963; Kostlan, 1954; Lindzey, 1965, Sines, 1959) have used as cues protocols from one or more projective tests, such as the Rorschach, Thematic Apperception Test (TAT), and Draw-A-Person Test (D-A-P). Mischel (1968) has devoted almost an entire chapter to a discussion of the low predictive validity of projective tests. In addition, the few replicable empirical relationships which do exist have failed to further the development of a theory of psychopathology. It is not surprising, therefore, that this type of information has demonstrated only marginal utility in clinical judgment situations.

A second example in clinical judgment research is the work of Goldberg (1965, 1969, 1970) who has conducted a very thorough program of research investigating relationships between MMPI profiles and diagnosis of psyshosis versus neurosis. In this case the source of difficulty is both the cue and the criterion. While the MMPI is a definite improvement over projective tests, the fact that it is far from ideal as an assessment device is well documented (Jackson & Messick, 1961, 1967; Rogers, 1971). The major problem is that an MMPI profile is not useful for the purpose of differential diagnosis because of the failure of the MMPI scales to demonstrate discriminant validity. In addition, the criterion of differentiating psychosis versus neurosis is not particularly useful. Whether considered as diagnostic categories or as a complex psychological dimension, this distinction is simply too gross.

It is common knowledge to those active in judgment research that attempts to train people to improve in judgments of the type described have failed after literally months of training. That this could happen should be no surprise considering the nature of the task.

This example was not chosen to single out one particular experimenter. To the contrary, it was chosen to demonstrate that even thorough and carefully executed clinical judgment research has been handicapped by a lack of meaningfulness in the judgment task, due to either criterion or cue inadequacies, a fact which Goldberg, himself, has acknowledged (1965, 1969).

Clinical Judgment and Consistency Issue

Situationists have cited widely the negative results reported in the clinical judgment literature as supportive of the view that future behaviours simply cannot be predicted from interview and test data. According to this view, transituational consistency is illusory, being constructed by the psychologist or lay observer. A program of research often cited in support of this view is the work of Chapman and Chapman (1971). This work deals with what the authors label "associatively based illusory correlation as a source of psychodiagnostic folklore." Basically, the reported series of studies demonstrates that in an experimental situation judges tend falsely to perceive co-occurrence of pairs of stimuli for which an associative or conceptual relationship exists. For example, if all possible pairs of a number of words were presented to a judge an equal number of times, the judge would tend to report that the pair salt-pepper appeared more often than the pair pepper-car. Similar findings are reported for pairs formed from symptom statement such as, "He is worried about how manly he is," and projective test responses, such as D-A-P drawings. That is, certain pairs which might appear to go together logically are judged to occur together when in fact there is no empirical basis for such a judgment.

These data sound a loud warning to those who continue to interpret invalid projective tests. However, they have little or no bearing on the issues of behavioural consistency and accuracy in trait inference. The fact that a subject will create consistency in an experimental situation where none is present has no relevance to the question of the existence of consistency. Again, it should be emphasized that the strong tendency for people to assume that consistency and order will be the rule, suggests that they have found this assumption to be valid in the past.

Most important, however, is that each of the cited research examples points out that a failure to insure the empirical validity, and hence psychological meaningfulness of an experimental task, can at best yield inconclusive results. At worst, such neglect may lead to false conclusions.

CHAPTER FOUR

A MODEL OF BEHAVIOURAL CONSISTENCY

Most research dealing with behavioural consistency or clinical judgment is concerned with one or both of two broad purposes. The first is some form of explanation of behaviour covariation, and the second is the prediction of future behaviour. These purposes require a well developed concept of behavioural consistency, which in turn relies on clearly defined units of behaviour. Both of these requirements, especially the latter, are absent in a large majority of clinical judgment and behavioural consistency studies.

This chapter attempts to satisfy these requirements with models of assessment and judgment based on clearly defined dimensions of behaviour, and a probabilistic model of consistency. Before presenting the model, it will be useful to review briefly the traditional model of consistency and its behavioural unit, the trait.

The Traditional View

Traits are most often defined as underlying dispositions to respond to a particular class of stimuli in a consistent manner. In addition, traits are usually considered causal and motivational agents. For example, Allport (1966, p. 1) viewed traits as having "...more than nominal existence" and as "...dynamic, or a least determinative, in behaviour." Cattell (1957) concurred with this causal and motivational view of traits. Mischel (1968, 1969, 1973), on the other hand, objects strongly to this view of traits as underlying attributes which <u>cause</u> behavioural consistency.

> Sometimes mental traits and states are invoked as if they were the causes of behaviour, while their own antecedents are ignored or forgotten. Unfinished causal sequences are found whenever mental states (cognitions, affects, motives, etc.) are employed as explanations while the determinants of the mental state themselves are omitted from the analysis. (Mischel, 1968, p. 95)

The observation of the circularity of citing dispositions as causes is sound. Concurring with Mischel's view, this causal concept was rejected in the formulation of a new model of behavioural consistency.

Behavioural Dimensions

A central feature of the model being proposed is that the use of the term trait is discontinued in favour of behavioural dimension. A behavioural dimension consists of a conceptual family or category of behaviours. For example, aggression refers to a category of behaviours which share certain characteristics. Physical attack, verbal abuse, and stepping in front of someone in a line, would each be considered aggressive behaviour. It is emphasized that while the relationship of these behaviours to aggression is conceptual in nature, one is nevertheless dealing with the relationships among behaviours, and not simply lexical relationships. This means obviously that empirical relationships can be assessed. Therefore, from an assessment point of view, depending on the purpose of the particular assessment or measurement enterprise, one will be concerned with either the frequency of occurrence or the probability of occurrence of some specific category of behaviours, as well as the relationships among various categories. From this point on these categories will be referred to as behavioural dimensions.

In this context then, behavioural consistency refers to an individual's probabilities of engaging in behaviours related to behavioural dimensions. The structural model for assessment scales is cumulative, meaning that the more items on the scale for which there are responses in the keyed direction, the greater is the probability of occurrence for behaviours relevant to the dimension under consideration. In a technical sense, a reference to a person as aggressive means simply that this individual has a relatively high base rate of aggressive behaviours. With respect to assessment, it would mean that the individual scored high on some measure of the construct aggression. The degree of behavioural consistency typically to be found can be empirically investigated in terms of the stability of these base rates and probabilities as assessed within various modes of measurement, across different situations. There are, of course, certain methodological problems which must be taken into account if one hopes to deal adequately with the question of behavioural consistency. Some of these, such as method variance and criterion assessment are discussed in later sections of this paper.

Item Responses as Samples of Behavioural Dimensions

A second major aspect of the new model has to do with the meaning of a response to an item, and the relationship between item responses and non-test behaviour. To develop a full understanding of these points, it is necessary first to review briefly certain aspects of construct validity.

In her classic monograph, Loevinger (1957) makes a distinction between traits and constructs. "Traits exist in people; constructs [in this case about traits] exist in the minds and magazines of psychologists" (p. 83). With this statement, Loevinger was trying to emphasize that the data are manifestations of traits existing within the individual, while constructs are psychologists' best understanding of these traits. In other words, constructs represent psychologists' best attempts at placing meaningful and valid conceptual frameworks around the manifestations of various traits. The form of behavioural manifestation dealt with for a particular trait is a function of the instrument being used or the particular mode of observation (Fiske, 1971) involved. In the case of structured psychological tests, item responses are the manifestations of the traits in question.

In the sense of manifestations of traits, Loevinger refers to item responses as signs of behaviour.[2] By this she means that the item in question and some potential criterion behaviour are both related to the same underlying trait-construct, which the test (of which the item is a part) is designed to measure. Therefore, a particular item response is an indication or sign of some other behaviour because they are both related to the same hypothetical construct.

[2] Loevinger also considers item responses as samples, but she is referring only to the fact that the act of responding to items is a behaviour, not to a relationship between item content and its behaviour referents.

The choice of the designation "sign" for this relationship seems unfortunate. This is especially so in the case of the cumulative structural model, which is utilized exclusively in the present approach. One reason the use of "sign" seems inappropriate is the connotation associated with the term with respect to projective techniques such as the Rorschach, and also with respect to instruments such as the MMPI, which were constructed using a blind empirical strategy. In both cases the relationship between a response and an external behaviour is rather mysterious and indirect. For many advocating the strictly empirical approach, this relationship is not even a matter of concern (Wiggins, 1973). The requirements of construct validity, however, are such that there should be a specific theoretical link between test and non-test behaviours based on the particular construct to which an item is hypothesized to relate. There is hopefully nothing mysterious about this relationship, especially when using a cumulative model. In such cases, the direct relationship between the item response and the external behaviour in question is a central feature contrasting it with alternative test construction strategies. In light of these considerations, it seems unfortunate that the phrase "signs of behaviour" was chosen.

Jackson (1971) previously discussed the advantages of considering items as samples of content. This idea is now extended in the proposition that items for a given scale are most profitably considered as a sample of behaviour, drawn within the mode of self report, and relevant to a specific behavioural dimension. Responses to items of this type constitute considerably more than just signs of behaviours in other modes. That is, because in this new model data collected within different modes still refer to the same common behavioural dimension, item responses and non-test behaviour both are considered samples from this single behavioural dimension.

Scale Scores as Measures of the Probability of Behaviour

In actual practice it is the relationship of the total scale score to non-test behaviours which is of interest. However, this degree of distance from the original data is no problem if certain criteria have been satisfied in the construction of the scale. First, the items of a scale should have substantive validity with respect to the behavioural dimension of which they are intended to be a sample (Jackson, 1970, 1971; Loevinger, 1957). Second, the items of a scale should be homogeneous as demonstrated by a high internal consistency reliability (Cronbach & Meehl, 1955, Jackson, 1970, 1971). Third, the items should be high in content saturation and written so as to suppress the effects of response styles (Jackson, 1970, 1971). Constructing scales which satisfy these criteria will provide total scale scores which are extremely reliable as compared with item responses, and which retain the property of being samples of their respective behavioural dimensions.

Scales constructed in the above manner also have the property that individual scores on the scale represent the relative probability of non-test behaviour relevant to the dimension sampled by the scale. Because a scale score represents the percentage of sampled behaviours which were endorsed in the mode of item responses, this score can be translated into a probability statement regarding the likelihood of the occurrence of other behaviours which are also samples of the dimension measured by the scale. As was mentioned above, it is the stability of an individual's relative probabilities for various dimensions of behaviour which is at the heart of the definition of behavioural consistency.

It should be noted that the probabilistic nature of behavioural consistency has important implications for the design of research. First of all, for any single discrete behaviour, the probability that it will occur is never precisely 1, nor precisely 0. In addition the fact that these relative probabilities clearly interact with situational factors (cf. Bowers, 1973) implies further that a single discrete behaviour should not be used as an external criterion behaviour. This assertion is consistent with results obtained by Fishbein and Ajzen (1974) in the context of predicting behaviour with measures of attitudes. Remember that properly constructed scales involve specific and at times elaborate procedures which begin as early as the creation of the original item pool (Jackson, 1970, 1971; Loevinger, 1957). To operate on the assumption that a single discrete behaviour is an adequate sampling of the behaviour dimension is to forfeit the game before it begins. One simply cannot test a theory of behavioural consistency using such a paradigm.

CHAPTER FIVE

MEASURING BEHAVIOURAL DIMENSIONS

Definitions have been developed and scales constructed for measuring a number of behavioural dimensions. Twenty scales of this type were developed for the Personality Research Form (Jackson, 1967, 1974). The dimensions assessed by this instrument focus on nineteen personality variables originally defined by Henry Murray (1938). These were modified where necessary to conform with research evidence, as well as with requirements of the models of behavioural consistency and scale construction underlying this approach to test development. One other personality variable is assessed which, along with two validity scales make a total of twenty-two scales. A notable aspect of the test construction strategy was the generation of items based on dimensions of item content which were rationally defined prior to any empirical investigations of item properties or scale reliability and validity (Jackson, 1967, 1970, 1971, 1974).

Empirical evaluation of PRF scales has provided evidence that the approach suggested here yields results superior to traditional approaches to personality assessment (Anastasi, 1976; Jackson, 1967, 1970, 1974; Skinner, Jackson, & Rampton, 1976). For example, median test-retest validity for the 20 content scales (Form AA) was .81, with a range of .69 to .90 (Bentler, 1964). The median internal consistency reliability (Kuder-Richardson formula 20) for the same form was .76 (Jackson, 1967, 1974), which is quite high for 20 item scales. Convergent validity of the scales also proved to be comparably high. The median validity correlation with behaviour ratings made by peers (college student subjects) was .52 (Jackson, 1967, 1974). This would appear to be a substantial increment over the .30 correlation cited by Mischel (1968) as the modal validity coefficient for personality research.

Probably the most important characteristic of the PRF is the successful suppression of response styles which facilitated the achievement of discriminant validity. The intercorrelations of the scales are modest, both with respect to the correlations among content scales and the correlations with desirability. What this means is that a number of independent dimensions of behaviour are being measured. Factor analyses have yielded either five orthogonal (Skinner, Jackson, & Rampton, 1976) or six (Stricker, 1974) to eight (Nesselroade & Baltes, 1972) oblique factors. This contrasts with instruments such as the MMPI, for which the first two factors account for a large portion of the total variance, primarily as a result of the high correlation of these factors with acquiescence and desirability responding (Jackson & Messick, 1961).

One recent study (Skinner et al., 1976) concerned with the stability of PRF scales and items when translated into French has provided strong evidence for the utility of the behavioural dimension approach. Items from the PRF-E were translated into French for use in placement research within the Canadian Armed Forces. The factor structures for 1040 French-speaking and 2141 English-speaking male volunteers were independently derived. The coefficient of congruence (Harman, 1967) which measures the cosine of the angle between factors equalled .99, .98, .98, .99, and .96 for respective factors of the English and French versions. This is an especially impressive degree of congruence when one considers that it is based on solutions independently rotated to a varimax criterion. That is, there were no procrustean rotations used to maximize congruence.

It is important to reemphasize the meaning of stable structures of covariation found among behavioural dimensions such as those measured by the PRF. What has been provided by analyses such as these is a mapping of the empirical relationships or covariations among behavioural dimensions. This mapping is essentially a structural model of potential patterns of behavioural consistency. What the model suggests is that a high relative probability that a given individual will engage in a particular category of behaviour, such as dominant behaviours, may imply a high relative probability exists also for some other category of behaviour, such as aggression. At the same time this same high probability for dominance may imply a low probability for an affiliative behaviour, and bear no particular relationship (i.e., suggest a .5 probability) to play behaviour. In other words, the highly stable structure among these behavioural dimensions suggests that one could identify individuals who display reliable patterns of behaviour probabilities. These patterns are typically presented graphically as a profile, where low points represent dimensions of behaviour for which the individual has a low probability of engaging, and high points represent dimensions of behaviour for which the individual has a high probability of engaging. Furthermore, because there exists a finite number of statistically independent behavioural dimensions, there also should be a finite number of statistically independent patterns of behaviour probabilities or profiles. This suggests the possibility of identifying groups of individuals sharing common patterns of behavioural probabilities.

Dimensions of People

A program of classification research has been investigating the proposal that groups of individuals exist sharing distinct patterns of behavioural probabilities (Skinner & Jackson, 1974; Skinner, Jackson, & Hoffmann, 1974; Skinner, Reed, & Jackson, 1976). This research differs from traditional classification research in that it is based on behavioural dimensions of the type described above. The primary classification instrument has been the Differential Personality Inventory, or DPI, which measures 27 behavioural dimensions relevant to psychopathology (Jackson & Carlson, 1973; Jackson & Messick, 1970), and was constructed using the same rationale as the PRF. The analytic strategy for classification, called Modal Profile Analysis, involves generalized canonical correlation analysis identifying dimensions of entities (persons) based on similarity in profile shape. A detailed presentation of the procedure has been given in Skinner (1975, 1976) and Skinner and Jackson (1974).

This methodology produces matrices of factor loadings and factor scores possessing interesting and useful properties. The factor loadings represent factors of entities or people. Individuals having high loadings (with the same sign) on a factor have similar patterns of scores on the set of scales or behavioural dimensions. The factor scores, on the other hand, describe the location or projection of each scale on the factors of the entity (people) space (Skinner, 1975, 1976). This means that if the Impulsivity scale has a high positive factor score for the entity Factor I, then impulsive behaviour has a high relative probability of occurrence for the people defining the positive pole of this factor (high positive loadings) and a low relative probability of occurrence for people defining the negative pole of the factor (high negative loadings). Each scale included in the analysis has a factor score for every factor. Thus, for each factor of people there is a characteristic pattern of scale scores, or relative behavioural probabilities, for the set of behavioural dimensions. A set of scale scores for a factor is called a Modal Profile.

This description should have made it clear that each modal profile actually represents two groups of people. That is, the factors produced by the modal profile analysis are bipolar factors. The modal profile, or set of factor scores, for a given factor is really the characteristic profile only for those subjects whose individual profiles had a high, positive correlation with the modal profile for that factor. These people are clustered around one pole (the positive pole) of the factor. The characteristic profile of those individuals clustered at the negative pole of the factor (high negative correlations with the factor scores) is obtained by simply reflecting the original factor score. Illustrated in Figure 1 is the effect this has on the profile, which is simply to reverse the order of saliency of the scales.

It should be noted that the complete procedure for Modal Profile Analysis actually involves a sequential strategy, requiring both within and between sample analyses. Details are provided in Skinner (1976). The results of a Modal Profile Analysis for 6 samples based on 25 DPI scales are shown in Tables 1a and b.

Using a criterion of a minimum loading of |.50|, 407 subjects, or 62.2 per cent, were classified to one of the Modal Profile factors. The high classification hit rate obtained for the 654 subjects indicates

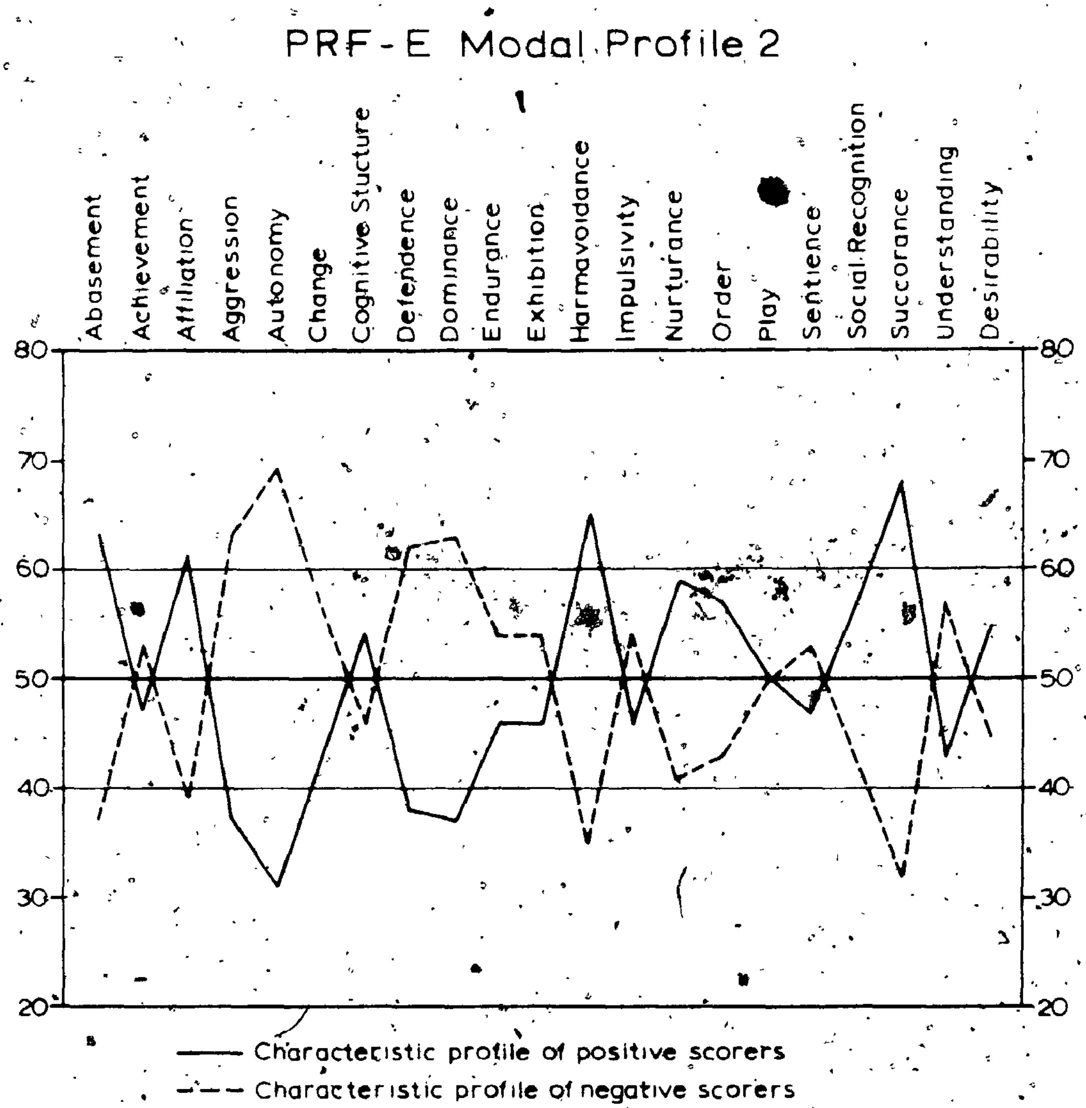


Figure 1. Characteristic Profiles for Positive and Negative Poles of PRF-E Modal Profile 2.

that the seven Modal Profiles are very robust. In interpreting this hit rate one should keep in mind that the Modal Profiles are derived from a sequential cross-validational design, and therefore this hit rate represents essentially a cross-validation hit rate. In fact, Skinner (1975) has recently reported cross-validation hit rates as high as 67%. The author is unaware of any other psychological classification research using an objective strategy which has even approached cross-validation hit rates of this magnitude.

Similar levels of success have been achieved with PRF data for college students and military personnel (Skinner, 1975). Five PRF Modal Profiles derived from six military and one college sample are reported in Tables 2a and b.

This series of DPI and PRF classification studies is taken as further evidence supporting the utility of employing clearly defined dimensions of behaviour in psychological research.

Table 1a

SALIENT SCALES

DIFFERENTIAL PERSONALITY INVENTORY MODAL PROFILES

(Derived from 3 Alcoholic, 1 Prison, and 2 Psychiatric Samples)

Modal Profile I+	Modal Profile I-
Defensiveness, Repression, Health Concern, Shallow Affect.	Depression, Familial Discord.
Modal Profile II+	Modal Profile II-
Rebelliousness, Shallow Affect, Socially Deviant Attitudes, Sadism.	Panic Reaction, Feelings of Unreality, Disorganization of Thinking, Health Concern, Perceptual Distortion.
Modal Profile III+	Modal Profile III-
Perceptual Distortion, Socially Deviant Attitudes, Health Concern, Hostility, Sadism.	Desocialization, Repression, Self Depreciation, Depression.
Modal Profile IV+	Modal Profile IV-
Impulsivity, Neurotic Disorganization.	Health Concern, Hypochondriasis, Ideas of Persecution, Familial Discord, Broodiness.
Modal Profile V+	Modal Profile V-
Sadism, Cynicism, Ideas of Persecution, Shallow Affect.	Rebelliousness, Somatic Complaints, Irritability, Hypochondriasis.
Modal Profile VI+	Modal Profile VI-
Irritability, Mood Fluctuation, Health Concern, Repression.	Hypochondriasis, Somatic Complaints, Perceptual Distortion, Shallow Affect.
Modal Profile VII+	Modal Profile VII-
Cynicism, Sadism, Somatic Complaints, Hypochondriasis.	Familial Discord, Defensiveness, Perceptual Distortion.

Table 1b

Multiprofile-Multisample Analysis of DPI Data for 3 Alcoholic, 1 Prison, and 2 Psychiatric Samples
(Total n=654)

	Broodiness	Cynicism	Depression	Desocialization	Disorg. of Thinking	Familial Discord	Feelings of Unreality	Health Concern	Hostility	Hypochondriasis	Ideas of Persecution	Impulsivity
Modal Profile 1	42	48	39	41	48	39	47	69	45	54	48	46
Modal Profile 2	45	56	49	57	38	55	36	39	57	47	45	50
Modal Profile 3	47	54	34	29	43	48	54	62	62	52	56	57
Modal Profile 4	39	40	46	48	54	38	53	37	49	37	37	80
Modal Profile 5	55	64	47	57	53	51	58	49	49	38	64	44
Modal Profile 6	59	53	51	59	41	45	41	63	59	28	47	51
Modal Profile 7	52	64	44	55	58	20	47	58	40	60	47	48

	Irritability	Mood Fluctuation	Neurotic Disorg.	Panic Reaction	Perceptual Distortion	Rebelliousness	Repression	Sadism	Self Depreciation	Shallow Affect	Soc. Dev. Attitudes	Somatic Complaints	Defensiveness
Modal Profile 1	44	45	45	46	56	46	73	50	42	62	46	53	77
Modal Profile 2	50	41	46	34	39	74	51	61	47	69	66	45	53
Modal Profile 3	53	57	50	47	65	56	33	60	33	41	64	49	43
Modal Profile 4	41	56	63	54	57	47	59	55	59	54	57	40	48
Modal Profile 5	37	42	44	46	59	29	44	72	54	61	49	34	49
Modal Profile 6	67	66	42	56	37	47	62	55	45	37	50	35	53
Modal Profile 7	56	41	59	48	37	50	53	64	54	45	56	62	35

Table 2a

SALIENT SCALES

PERSONALITY RESEARCH FORM E MODAL PROFILES

(Derived from 1 College and 6 Military Samples)

Modal Profile I+	Modal Profile I-
Understanding, Achievement, Endurance, Nurturance, Desirability.	Harmavoidance, Defendence, Impulsivity, Aggression, Play.
Modal Profile II+	**Modal Profile II-**
Succorance, Harmavoidance, Abasement, Affiliation.	Autonomy, Aggression, Dominance, Defendence.
Modal Profile III+	**Modal Profile III-**
Exhibition, Play, Affiliation, Impulsivity, Nurturance.	Cognitive Structure, Harmavoidance, Order, Defendence.
Modal Profile IV+	**Modal Profile IV-**
Dominance, Social Recognition, Exhibition, Defendence.	Abasement, Autonomy, Impulsivity.
Modal Profile V+	**Modal Profile V-**
Exhibition, Change, Play, Desirability, Affiliation.	Sentience, Succorance, Nurturance.

Table 2b

Multiprofile-Multisample Analysis of PRF-E Data for One College and Six Military Samples

(Total n=796)

	Abasement	Achievement	Affiliation	Aggression	Autonomy	Change	Cognitive Structure	Defendence	Dominance	Endurance	Exhibition	Harmavoidance	Impulsivity	Nurturance	Order	Play	Sentience	Social Recognition	Succorance	Understanding	Desirability
Modal Profile 1	55	62	53	36	47	53	52	35	59	62	52	34	35	60	56	37	57	41	42	63	60
Modal Profile 2	63	47	61	37	31	43	54	38	37	46	46	65	46	59	57	50	47	57	68	43	55
Modal Profile 3	55	43	63	47	46	55	34	40	53	45	68	34	61	60	37	66	57	48	51	42	45
Modal Profile 4	28	51	56	55	29	41	56	60	66	49	63	45	40	51	49	44	50	64	56	45	51
Modal Profile 5	43	48	61	41	58	62	55	43	54	46	64	59	41	38	59	62	29	47	37	41	62

CHAPTER SIX

PERCEIVING BEHAVIOURAL DIMENSIONS

The empirical validity and utility of behavioural dimensions are relevant to the controversy concerning the accuracy of trait inference and implicit personality theory, the subject of Chapter Two. To review briefly, Lay and Jackson (1969) reported marked similarity in structure between perceived behavioural co-occurrence and the structure of actual behavioural co-occurrence as assessed by factor analysis of PRF scale scores. The controversy centers on the authors' conclusion that the layman's view of the network of implicative relations among behaviours is for the most part veridical. The high reliability and demonstrated convergent and discriminant validity of PRF scales, as well as the successful replication of PRF factor structure in several investigations, should serve to eliminate any lingering doubts that the PRF factor structure bears a relationship to dimensions of covariation or co-occurrence of non-test behaviour.

However, the answer to the question of the veridicality of perceived behaviour covariation requires more than just the demonstration of similar structure for judged and empirical co-occurrence of behaviours combined with the fact that the measures of empirical co-occurrence are valid. One must also be concerned with the reasons that behaviour is perceived to covary in the same fashion that it actually does. One explanation which has been advanced is that perceptual biases are responsible for congruence of perceived and actual behaviour covariation. In response to this argument, early sections of this paper dealing with trait inference and clinical judgment pointed out that with the exception of judgment situations which are contrived or ecologically invalid, biases are not a major source of the reported congruence.

The alternative explanation being advocated here is that a major component of the congruence in structure between perceived and empirical behaviour covariation is due to veridical perception. Of course, the possibility remains that while an aggregation of individuals may perceive correctly the general structure of behaviour covariation, specific inferences about individual stimulus persons by individuals or groups may be considerably less accurate.

Behavioural Dimensions in Person Perception and Clinical Judgment Research

A primary concern of person perception and clinical judgment research is the investigation of the accuracy of the individual judge in making specific inferences regarding a stimulus person. Research in these areas should be facilitated by the availability of an empirically validated network of behaviour covariation, based on clearly defined dimensions of behaviour. Recall the repeated instances cited above where difficulties in clinical judgment research were associated with a failure to relate judgment tasks to empirically valid relationships. Recall also the problems caused by poorly defined or inadequately measured criteria. It is asserted that clinical judgment and person perception research in which judgment tasks are conceived in terms of behaviours sampled from well defined dimensions with established empirical relationships, will yield results differing significantly from previous efforts. Indeed, studies have already been published supporting this view (Lay, 1970, Lay, Burron, & Jackson, 1973; Lay & Jackson, 1969, Reed & Jackson, 1975).

In particular, the study of the processes underlying inferential accuracy (Jackson, 1972) in clinical judgment and person perception should be affected. Inferential accuracy refers to correct prediction or inference regarding the behavioural probabilities of a stimulus person, based on a limited sample of behaviour of that same stimulus person. A model of the process of inferential accuracy (Jackson, 1972) is currently under investigation (Reed & Jackson, 1975). This research program should help fill in gaps in what is known about accuracy of the individual perceiver, and the processes underlying inferences regarding the co-occurrence of behaviour in others.

A Model for Inferential Accuracy

Inferential accuracy is defined in terms of a person's ability, given limited information about the past behaviour of a target person, to judge correctly the probabilities of other behaviour by that person, and to identify behavioural exemplars as part of a pattern of behavioural consistencies. This implies that in making judgments about the personalities of other people, judges rely to a large extent upon expectations they hold regarding the covariation of behaviours in others. A similar assertion was made first by Asch (1946) with respect to trait co-occurrence. Thus, a judge, knowing that a person is irritable, might infer behaviour relevant to hostility, but not necessarily to hypochondriasis. Jackson's model of inference postulates that two distinct processes underlie the use of one's knowledge of the network of behaviour covariation in making judgments about others, and additionally that individuals will vary with respect to these processes. Specifically, it is suggested that given limited information, individuals vary in terms of their awareness or sensitivity to the network of behavioural consistencies, and in terms of their readiness or threshold to attribute behaviours to others based on the implicative relations among behaviours.

An example should help clarify the proposed model. Consider the following description of a mental patient, taken from Reed and Jackson (1975):

> Jack Cole has been arrested several times for theft. Usually his crimes have been poorly planned and rather reckless. He says that he does not feel guilty about his behaviour and often explains his stealing something by simply saying he wanted it. In interviews Jack frequently mentions his strong dislike for rules and discipline, and he seldom speaks of friends.

This description actually was based on Modal Profile II+ from Skinner et al (1974). Now consider the following behaviour exemplars. Each is a positive (keyed) exemplar of the dimension to its immediate left.

(Rebelliousness)	I like to do the opposite of what other people do.
(Impulsivity)	I'm willing to do almost anything on the spur of the moment.
(Depression)	My present sitatuion seems quite hopeless.
(Defensiveness)	I cannot think of any way in which I have failed a friend.
(Hypochondriasis)	I can never really locate my illness.

Assume that 200 judges have rated the probability that Jack Cole would respond "True" to each of the five statements on a scale ranging from 9, extremely likely to answer "True", to 1, extremely unlikely to answer "True". Averaged over 200 judges, the ratings for five dimensions are as follows: Rebellious - 9.0, Impulsive - 7.0, Depressed - 5.0, Defensive - 3.0, Hypochondriacal - 1.0, where the average value represents the layman's conceptions regarding the ordering of behavioural probabilities for this target person. For the present, assume that the group consensus of the target's behavioural probabilities is accurate. This assumption seems justified in light of the lengthy arguments presented above. However, it lacks extensive empirical validation, a situation which should be remedied by the research reported in Chapter Eight.

Consider Judge A, who rated the five behavioural probabilities 9.0, 7.0, 5.0, 3.0, and 1.0, mirroring the group consensus exactly, and thus is highly sensitive to the consensus, while Judge B, who rated the five behavioural probabilities 3.0, 3.0, 3.0, 3.0, and 3.0 has patterned his ratings independently of the group consensus, thus providing evidence of lack of sensitivity (see Figure 2a). A judge's sensitivity can be estimated by the product-moment correlation between his or her individual judgments and group consensus values.

Threshold may be conceptualized as the psychological point at which the judge begins to ascribe positive relationships between a target and a behaviour. That is, threshold is the point at which a judge feels that the response, "True," is more likely than, "False." This has been illustrated graphically in Figure 2b by dropping a perpendicular to the group consensus axis, from the judgment value representing a point where the judge is clearly ascribing endorsement of the behaviour exemplar to the target (5.2 in Figure 2b). The point at which the perpendicular intercepts the group consensus might be taken as the threshold estimate (Jackson, 1972). This

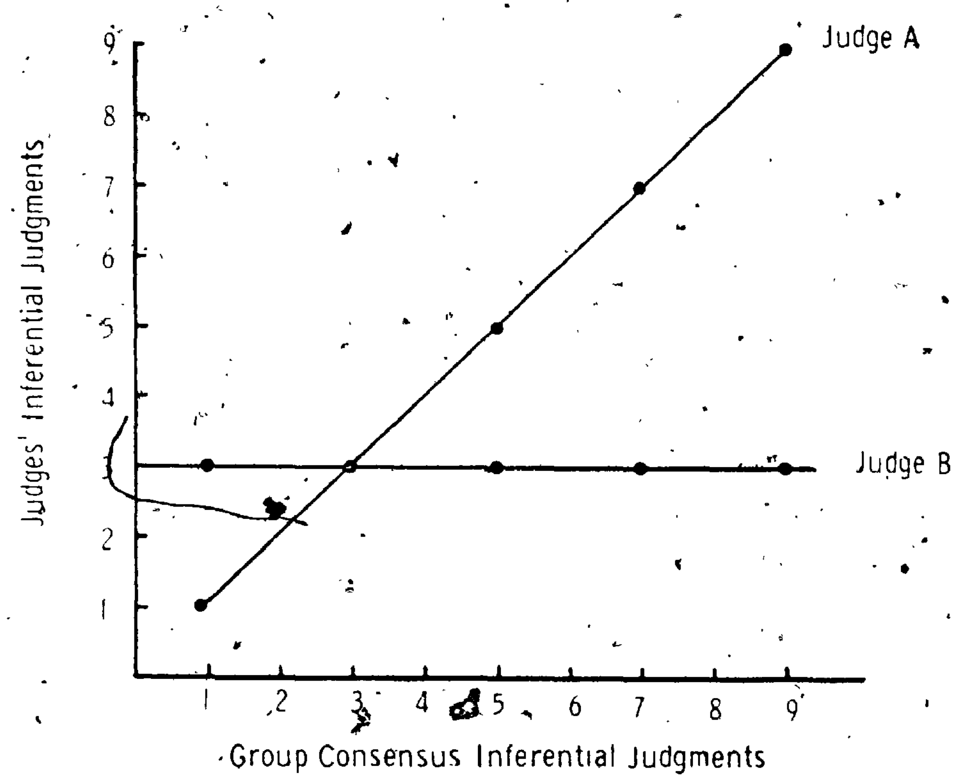


Figure 2a. Inferential Judgments of Two Hypothetical Judges Differing in Sensitivity.

method of estimating threshold is for the most part quite straightforward both with respect to the psychological meaning of threshold and the mathematical properties of the estimate. However, one problem does exist, namely that it is mathematically possible for threshold to assume values less than zero. A judge's mean judgment for a target has the same psychological meaning as the earlier threshold esimate, while avoiding this one difficulty. Thus, the mean has been adopted for use as the superior threshold estimate. Because low mean and high threshold both indicate a reluctance to attribute behaviours, the mean value must be reflected on the nine-point scale so that the magnitude of the resulting threshold estimate will be consistent with the definition of threshold.

To illustrate, compare Judge C, who rated the five traits in the above example 2.0, 3.0, 4.0, 5.0, and 6.0, with Judge D, who rated the five traits 5.0, 6.0, 7.0, 8.0, and 9.0 (Figure 2b). Although both accurately reflect the rank and patterning of the group consensus values, Judge D shows a much greater willingness to ascribe relatively remote behaviours as characteristic of the target than does Judge C. Hence, D's threshold is lower as evidenced by his willingness to describe the target as likely to engage in behaviours which others consider less probable. Judge C, on the other hand, uses much more stringent criteria.

In summary, sensitivity can be thought of as an individual's awareness of and use of the consensus regarding the covariation of behaviours. Threshold, on the other hand, represents an individual's willingness to attribute behaviours to a target, based on implicative relationships among the behaviours.

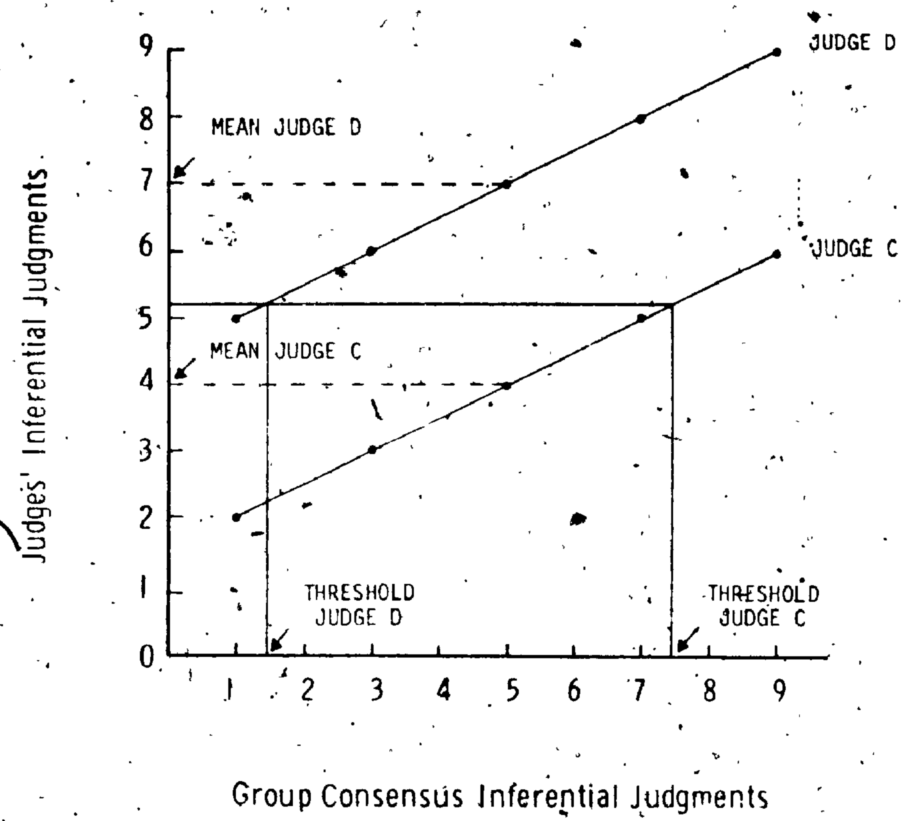


Figure 2b. Inferential Judgments of Two Hypothetical Judges Differing in Threshold.

Preliminary Empirical Investigation of Inferential Judgments

A study has been completed examining certain properties of the model for inferential accuracy (Reed & Jackson, 1975). Presented with three targets based on the original Skinner et al. (1974) modal profile analysis, and one target representing a normal well adjusted invididual, 189 subjects judged the likelihood that each target would endorse 54 Differential Personality Inventory items. These items represented one true-keyed and one false-keyed behavioural exemplar for each of the 27 dimensions assessed by the DPI. The target descriptions for the three clinical targets contained behaviours characteristic of those dimensions for which representative patients had reported the highest behavioural probabilities. That is, they represented the four or five scales having the highest score on the associated modal profile. The descriptions are reproduced in Table 3.

Reliability of the Group Consensus

The split half reliability of judges' consensual judgment for each target was determined by first dividing the subjects into two randomly-selected groups (n = 94 and n = 95), calculating independent group consensus values for each group, and then computing the correlation between the two sets of group consensus values. This was carried out for each of the three clinical targets, as well as for the non-clinical target. The resulting correlations were all in excess of .94. When corrected by the Spearman-Brown formula to estimate the reliability based on all 189 judges, reliabilities were equal to .97, .99, .99, and .99, for targets I through IV respectively. Thus, the group consensus values for all targets were extremely

Table 3

Target Descriptions from Reed and Jackson (1975)

Target I

Stu Johnson has been married ten years, has two children, and has a good job in an office. In spite of these accomplishments he recently has been feeling worthless and says he doesn't deserve what he has. He feels tired all the time, and complains that he generally feels "pretty low." Stu has told his friends that he feels downhearted. They have been sympathetic, but he doubts that they mean it. He often says to his wife that if he ever really needed friends they would probably be too busy. (Clinical Depression: Cynicism, Depression, Self Depreciation, Somatic Complaints).

Target II

Jack Cole has been arrested several times for theft. Usually his crimes have been poorly planned and rather reckless. He says he does not feel guilty about his behaviour and often explains his stealing something by simply saying that he wanted it. In interviews Jack frequently mentions his strong dislike for rules and discipline, and he seldom speaks of friends. (Psychopathy: Desocialization, Impulsivity, Rebelliousness, Socially Deviant Attitudes).

Target III

John Bailey is a very quiet person. He keeps his thoughts to himself and is extremely reluctant to share his ideas with others. This is due to a general mistrust of people and a lack of faith in their motives. Although he seems to love his wife, there are many unresolved conflicts in their marriage. John simply will not discuss anything. He spends much time by himself thinking and often brooding over his problems. He has become obsessed with the idea that his wife is "cheating" on him and is planning to leave him for another man. (Pre-paranoid: Broodiness, Cynicism, Ideas of Persecution).

Target IV

Art Reynolds is quite happy with his life. He is married, has two children, and has a successful business of his own. He finds his work to be interesting and challenging; but he is still able to find time for his family and other interests as well. While Art, of course, has some problems and disappointments, most of his friends agree that things are going well for him. (Normal).

stable, implying that judges agreed very well concerning which specific behaviours were associated with particular modal behavioural patterns. This in turn provided a firm foundation on which to base the measurement of sensitivity.

Independence and Generalizability of Sensitivity and Threshold

Four sets of sensitivities and thresholds, one set each for targets I - IV, were calculated for each judge. Sensitivity was estimated by the correlation between an individual's judgments and the group consensus, and threshold by the reflected mean judgment. One would expect the two processes, defined in terms of mean and correlation, to be mutually independent. This proved to be the case as evidenced by the low correlations between the two. Correlations between sensitivity and threshold were -.07, -.08, .12, and .02, for targets I - IV respectively.

Generalizability estimates are .65 ($p < .001$) for sensitivity and .64 ($p < .001$) for threshold. These are based on the average of the correlations between scores (Nunnally, 1967, Eq. 6-18) for the three clinical targets only. Being determined from essentially only three items (e.g., each subject has one sensitivity score for each of the three clinical targets), these represent quite substantial generalizability indices.

Accuracy of the Group Consensus

A second question regarding the consensual judgment is whether it constitutes an accurate perception of the target described in the paragraph. In other words, do judges give high ratings to items keyed on scales which were the salient dimensions of the modal profile for which a particular target description was developed. To determine this, it was first necessary to reflect the mean ratings for false-keyed items, and then to sum, by scale, the mean ratings for the items keyed on each scale. The result of these calculations, reported in Table 4, reflect the considerable degree of accuracy present in the consensual judgments. For instance, of the eleven marker scales, nine received highest scores on the target for which they were markers (Cynicism is counted as a hit for Target III and a miss for Target I). One of the two exceptions, Desocialization (marker for Target II: Psychopathy) received its highest scores on Target III (Pre-paranoid), which, while not written specifically to include this trait, certainly did contain content emphasizing behaviours relevant to it. In addition, the score (14.0) for Desocialization on Target II is considerably higher than the mean score for that target (11.1), and ranked fifth of all scales for Target II, exceeded only by the other markers, Rebelliousness, Socially Deviant Attitudes, and Impulsivity, and by one nonmarker scale, Hostility (14.3, ranked 4th).

The rank orders of scale scores within each target provide additional support for the presence of accuracy in the group consensus. The mean within target rank for marker scales was 4.9, as compared to 15.4 for the nonmarker scales. The ranks of the scale scores for the markers of each target were: Target I - 3, 6, and 13; Target II - 1, 2, 3, and 5; and Target III - 3, 4, and 6. An explanation for the difficulty in judging Target I is suggested below.

Also reported in Table 4 are the group consensus scale scores for Target IV, the non-clinical description. At least two aspects of the Target IV group consensus are noteworthy. First, this target

Table 4

Scale Sums of Group Mean Ratings Across Four Targets

(from Reed & Jackson, 1975)

Scale	Target 1 Clinical Depression	2 Psychopathy	3 Pre-paranoid	4 Normal
Insomnia	14.3	8.5	14.7	4.4
Headache Proneness	15.0	8.5	13.0	7.4
Broodiness	11.1	12.3	14.6**	5.9
Cynicism	12.2**	13.9	14.4**	6.6
Depression	14.2**	9.8	13.6	3.7
Desocialization	10.3	14.0**	15.2	4.1
Disorganization of Thinking	8.9	8.1	8.1	5.7
Familial Discord	9.9	14.0	12.3	5.1
Feelings of Unreality	13.3	11.3	13.5	6.6
Health Concern	10.7	9.3	10.6	10.4
Hostility	10.3	14.3	8.5	9.0
Hypochondriasis	10.1	9.9	12.1	7.4
Ideas of Persecution	11.8	13.0	13.8**	4.8
Impulsivity	8.2	15.5**	8.0	8.2
Irritability	13.3	13.4	14.2	7.4
Mood Fluctuation	11.2	8.5	10.8	10.1
Neurotic Disorganization	9.1	11.6	9.2	5.9
Panic Reaction	11.2	7.0	11.2	6.1
Perceptual Distortion	8.9	8.6	10.1	6.4
Rebelliousness	9.1	16.4**	9.4	8.4
Repression	11.3	10.5	11.7	7.8
Sadism	9.0	13.4	8.3	8.3
Self Depreciation	11.0**	7.8	8.0	6.4
Shallow Affect	9.6	7.8	8.8	7.3
Socially Deviant Attitudes	9.1	15.6**	7.4	8.1
Somatic Complaints	11.7**	9.6	10.4	7.8
Defensiveness	9.4	6.5	12.2	12.6
Target Mean	10.9	11.1	11.3	7.7

*Each scale score is the sum of the group consensus ratings for one true-keyed and one false-keyed (reflected) item.

**Marker scales for clinical targets.

received the highest score on Defensiveness of all four descriptions. Apparently, the judges felt that Target IV was "holding back," or that the picture painted by this description was a little too good to be true. Second, inspection of Table 4 shows that the mean scores on almost all of the remaining 26 scales were substantially lower for Target IV than for the three clinical targets. This reflects judges' correct attribution of greater psychopathology to clinical targets as opposed to non-clinical. This conclusion was supported by a one-way analysis of variance in which target was considered the independent variable, mean scale score the dependent variable, with 27 observations (the 27 scales) for each of the four target conditions ($F = 18.3$, $df = 3/78$, $p < .001$).

Table 5 shows the intercorrelation (based on 54 items) among group consensus, including that for judgments of the frequency of endorsement for "the average college student." Consistent with findings reported above pertaining to the relationships between the perception of clinical versus non-clinical targets, one finds that the consensus for the non-clinical or "normal" target has a moderately negative correlation ($r = -.44$) with each clinical group consensus. At the same time, one finds a very high and positive relationship ($r = .80$) between Target IV (non-clinical) and judgments of the frequency of endorsement for the "average college student." This pattern of correlations lends support to the assumption that Target IV represents a relatively "normal" individual, and also corroborates the findings reported above regarding judges' correct attribution of greater psychopathology to clinical as opposed to non-clinical targets.

Table 5

Correlations Among Four Group Consensus Vectors

and Judged Frequency of Endorsement

(Based on 54 items)*

	1	2	3	4	5
Target Person 1 (Depression)	1.0				
Target Person 2 (Psychopathy)	-.04	1.0			
Target Person 3 (Pre-paranoid)	.72	.13	1.0		
Target Person 4 (Non-clinical)	-.47	-.31	-.53	1.0	
Judged Frequency 5 of Endorsement	-.09	-.05	-.27	.80	1.0

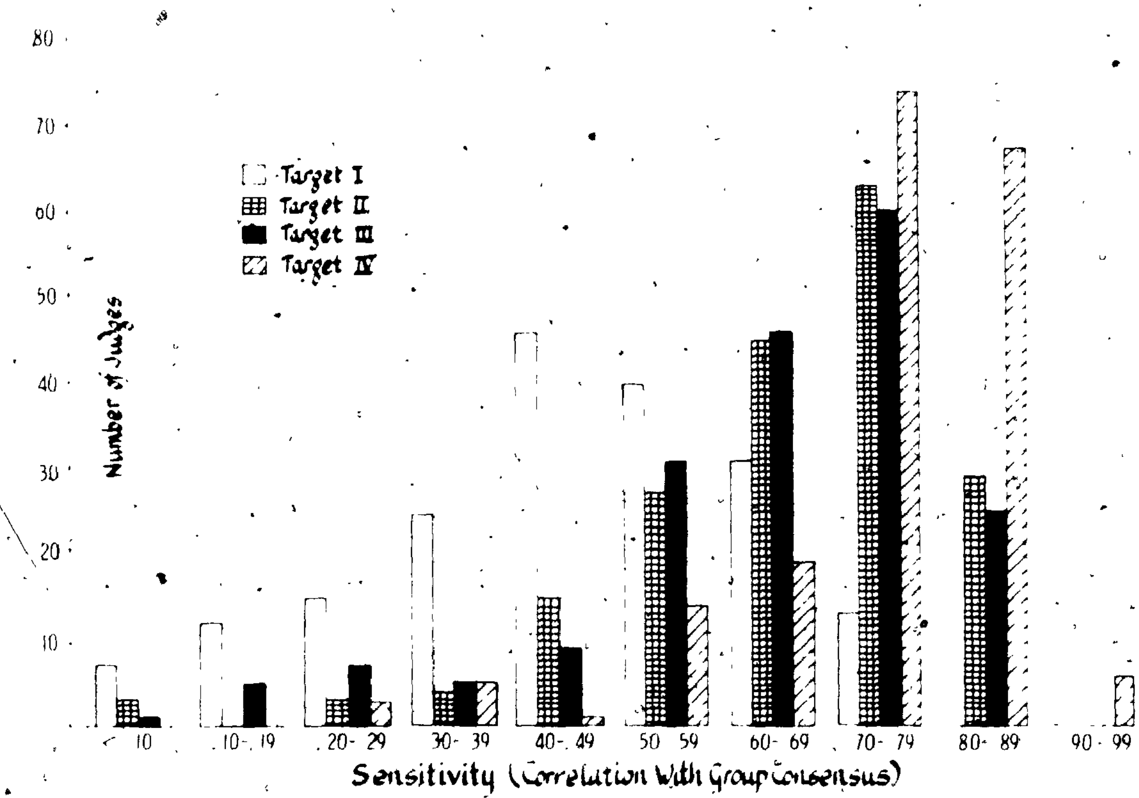
*For 52 df, $p < .01$ that $|r| \geq .35$

It should be emphasized that while the average judgment of the group is highly stable there is a distribution of individual sensitivities to this consensual judgment. Figure 3 presents frequency distributions of sensitivity for each target. Similar distributions were present for threshold as well. Evident from Figure 3 is the lower agreement among judges concerning Target I, than is the case for the other targets. Recall also the apparently lower accuracy reported for Target I. These data might be explained by the lower representation in the general population of this target relative to the others (Skinner & Jackson, 1974; Skinner, Reed, & Jackson, 1976). The lower base rate of occurrence could well have made the inferential task more difficult for Target I than it was for the other targets.

One should also observe in Figure 3 the extremely similar distribution of sensitivities for Targets II and III. What is noteworthy about this observation is that the extremely similar distributions and means (which are reasonably high) derive from criteria that are quite distinct. In fact, the correlation between the group consensus values for Targets II and III was only .13 (see Table 5).

A Need for Further Investigation of Group Consensus Accuracy

These findings are taken as quite encouraging. One observes fairly consistent levels of ability in inferring behaviours with regard to quite distinct targets as evidenced by the generalizability estimate for sensitivity. One also finds that the group consensus is extremely reliable, and that evidence is present supporting


80
70
60
50
40
30
20
10
Target I
Target II
Target III
Target IV
Number of Judges
< .10
.10-.19
.20-.29
.30-.39
.40-.49
.50-.59
.60-.69
.70-.79
.80-.89
.90-.99
Sensitivity (Correlation With Group Consensus)

the accuracy of the group consensus judgment. This implies that individual sensitivity, being assessed with reference to the group consensus, reflects a measurement of the individual judge's level of accuracy in making behavioural inferences.

However, as mentioned previously, there has been no direct empirical validation of the group consensus judgment. While it is anticipated that the group consensus for various targets will prove to be quite accurate, undoubtedly there will be some nontrivial discrepancy between group consensus and a more objective criterion. If one wishes to test the relationships such as those between individual sensitivity and other socially significant variables, then one must use the best available criterion to assess sensitivity. One of the primary aims of the research described in Chapter Eight is to carefully examine potential criteria for assessing sensitivity.

CHAPTER SEVEN

REVISIONS IN THE MODEL FOR INFERENTIAL ACCURACY

Experience with conducting research in the area of inferential accuracy has led the author to incorporate certain changes into the original Jackson (1972) model. The changes are concerned primarily with measurement and analysis procedures, and do little to change the nature of the psychological processes which Jackson suggested to be underlying the trait inference judgment. The modifications fall into two major categories: (1) level of analysis; and (2) parameters of the model. These modifications and the reasons for them are described below.

Level of Analysis Changes to Scale Scores

As explained in Chapter Five, a modal profile is derived from analysis of subject scale scores which identify groups of individuals sharing common patterns of behavioural consistency. Also explained above was the fact that each target description is based on the salient scales of a particular Modal profile. It seems desirable, therefore, that the assessment of accuracy of inferences about a particular target's behaviour be conducted also at the level of scale scores. While the assessment of accuracy considers scale scores or dimensions of behaviour, the individual judgments of a subject will still deal with the probabilities of occurrence for discrete behaviours. Thus, one has the best of both worlds: (1) judgments maintain the realism inherent in dealing with single discrete behavioural exemplars; while (2) one achieves the increased measurement fidelity gained by analyzing scale scores rather than individual items. It is for these reasons that the level of analysis has been changed from items to scale scores.

Model Parameters

A number of additional advantages derive from the use of scale scores in analyzing inferential accuracy. These concern the mathematical and psychological meaning of the parameters of the model.

Elevation. Threshold now becomes simply a linear transformation of profile elevation. In addition to the mathematical advantages described in a later section, this may have considerable psychological importance as well. For example, in the case of judgments of DPI targets, elevation is directly related to a judge's tendency to over- or underestimate a target's degree of psychopathology, as measured by the test. Because elevation can be computed more directly than threshold, and for a number of additional reasons which will become evident in the next two sections, this parameter will now be called elevation.

Sensitivity. Sensitivity is now equivalent to similarity in profile shape. Profile shape is a commonly understood concept in assessment. Thus, the meaning of sensitivity is more easily communicated. Sensitivity is simply the correlation between the profile of the subject's judgments of a particular target and the appropriate criterion profile for that target. Of course, the selection of the criterion profile is quite important, and is the concern of the investigations to be described in the next chapter.

Dispersion. A third parameter, dispersion, is introduced to the model. Dispersion is defined as the variance of scale scores about the subject's mean or elevation. This third parameter, thought to be primarily a response style reflecting one's use of the extreme categories of the rating scale, makes possible a more complete development of the components of inferential accuracy. However, before this is presented, it is necessary to discuss an earlier effort to isolate components of accuracy.

Isolating the Components of Accuracy

In 1955, Cronbach published a paper which has stopped nearly all research dealing with accuracy in various types of interpersonal perception. Cronbach pointed out that the studies up to that time had been based on global scores of accuracy. He demonstrated that such global scores confounded several mathematical components which had little to do with what most individuals would consider relevant to judgmental accuracy. Cronbach presented a detailed mathematical development of components of global accuracy scores. Few studies in the psychological literature have had the impact on a given area of research that this series of papers by Cronbach has had on the area of judgmental accuracy. First of all, it showed that the results of previous studies were uninterpretable. Secondly, except for a few rather isolated studies (i.e., Cline & Richards, 1960; Crow & Hammond, 1956; Sechrest & Jackson, 1961), accuracy research ceased, and was replaced by an emphasis on the process of judgment (Hastorf, Schneider, & Polefka, 1970; Tagiuri, 1969).

Cronbach's excellent papers were badly needed in this rather confused area of research. Unfortunately, the abrupt introduction of a high level of sophistication in the required analytical procedures sent accuracy research into a state of shock. Most researchers simply abandoned the area, and the results of those who did pursue it using the analytical procedures suggested by Cronbach (Crow & Hammond, 1957; Cline & Richards, 1960, 1961) were discouraging in their inconclusiveness. In addition, the four components of accuracy which Cronbach isolated (elevation, differential elevation, stereotype accuracy, and differential accuracy) were entirely abstract, and offered little with respect to increasing our understanding of the psychological processes underlying accuracy in interpersonal judgment. In my opinion, this is the primary reason for the abandonment of accuracy research. An individual interested in this type of research could not argue with Cronbach's mathematics. However, the components which his analysis produced were not interesting to many researchers because of their fundamentally mathematical as opposed to psychological nature.

The present author has argued (Reed & Jackson, 1975), as have others (Green, 1968; Jackson, 1972), that it is possible to develop components of judgmental accuracy which are both psychologically meaningful and mathematically respectable. In conjunction with this argument a number of problems in the interpretation of Cronbach's components need be discussed. Following this, when the sources of difficulty in this early analysis of accuracy scores are clearly understood, the model for inferential accuracy will be more fully developed and its advantages pointed out.

In order to compare the two approaches, it will first be necessary to briefly present the components of accuracy as developed by Cronbach. Wiggins (1973) has put together a very readable summary of Cronbach's original formal mathematical presentation of the components of the global accuracy scores. The following presentation borrows heavily from Wiggins' contribution.

The global accuracy score which was typically used prior to the 1955 paper was the average of squared differences between the n predicted and actual responses of k others (targets):

$$\frac{1}{kn} \sum_t \sum_i (y_{ti} - x_{ti})^2 ,$$

accuracy correlations. In general principle, I certainly support the concept of unconfounding distinct components and removing biases from accuracy correlations. However, removing from an accuracy score what Cronbach calls stereotype accuracy could in some instances remove the most meaningful or important component of accuracy. This would occur, for example, when the set of targets is quite homogeneous in some respect, such as when all targets are psychiatric patients. In this case, the mean of the targets would represent a general factor of psychopathology, and thus would represent an extremely important and meaningful component of the interpersonal judgment task. In this instance one might want to consider this component separately, but one would not want to arbitrarily remove it.

From the above example one can see that the removal of SA from the accuracy score may not produce the desired results. Cronbach was profound in pointing out the necessity of making distinct the concepts of sensitivity to the population mean and sensitivity to the behaviour of a single individual. The problem is that, in most cases, his formula for SA will not accomplish this. By using behaviour exemplars for which the population response frequencies are known, the model for inferential accuracy is easily able to accomplish what the removal of SA was intended to do. It is a simple matter to compute a profile of behavioural probabilities representing the population means for behavioural dimensions. This population profile can then be used to statistically remove from a judge's sensitivity score the effects of high sensitivity to the population frequencies of behaviours. In order to see more easily how this is accomplished it is necessary to look more closely at the model for inferential accuracy.

Figure 4

Expression of Global Accuracy Score as the Sum of Four Components

(After Cronbach, 1955; and Wiggins, 1973)

Component	Component Name
$\frac{1}{kn}\sum_t \sum_i (y_{ti} - x_{ti})^2$	Overall Accuracy
$= (\bar{y}_{..} - \bar{x}_{..})^2$	Elevation (E)
$+ \frac{1}{n}\sum_t [(\bar{y}_{t.} - \bar{y}_{..}) - (\bar{x}_{t.} - \bar{x}_{..})]^2$	Differential Elevation (DE)
$+ \frac{1}{k}\sum_i [(\bar{y}_{.i} - \bar{y}_{..}) - (\bar{x}_{.i} - \bar{x}_{..})]^2$	Stereotype Accuracy (SA)
$+ \frac{1}{kn}\sum_t \sum_i (y'_{ti} - x'_{ti})^2$ [a]	Differential Accuracy (DA)

[a]

$y'_{ti} = y_{ti} - \bar{y}_{t.} - \bar{y}_{.i} + \bar{y}_{..}$

$x'_{ti} = x_{ti} - \bar{x}_{t.} - \bar{x}_{.i} + \bar{x}_{..}$

have the greatest elevation.

Stereotype Accuracy (SA). Stereotype accuracy indicates ability to judge the average response of all the targets. This score depends on a judge's knowledge of the relative frequency of possible responses (Cronbach, 1955). This component may also be expressed as the variance of a difference:

$$SA^2 = \sigma^2_{\bar{y}_{.i}} + \sigma^2_{\bar{x}_{.i}} - 2\sigma_{\bar{y}_{.i}}\sigma_{\bar{x}_{.i}} r_{\bar{y}_{.i}\bar{x}_{.i}}$$

The correlation term reflects a judge's ability to predict the item means.

Differential Accuracy (DA). The differential accuracy term is concerned with a judge's ability to predict differences among targets for any item. Expressed as the variance of a difference, this component becomes

$$DA^2 = \Sigma[\sigma^2_{y'_{ti}} + \sigma^2_{x'_{ti}} - 2\sigma_{y'_{ti}}\sigma_{x'_{ti}} r_{y'_{ti}x'_{ti}}]$$

As the above expression implies, one can obtain a separate score for each item, or a total score over all items. There is a correlation term for each item representing the judge's ability to predict which targets will have the highest score on the item when the scores are expressed as deviations from a target's own mean.

This completes the description of the four components. As Cronbach (1955) and Wiggins (1973) have pointed out, the correlation terms of the DE and DA components are the most interesting and meaningful as measures of accuracy in interpersonal perception. The DE correlation term reflects sensitivity to differences among targets in overall elevation. The DA correlation term for a given item indicates a judge's sensitivity to differences among targets in the extent to which a target's item score is high or low relative to its own mean for all items.

Understanding this analysis of global accuracy scores is not an easy task. One reason for the difficulty is the rather large number of contingencies Cronbach attempts to account for simultaneously. For example, he included both accuracy in predicting for a single target, and accuracy in differentiating among different targets. In the case of DA these two different types of accuracy are actually confounded in the same component. Recall that while the DA component is concerned with differences in an item's scores among targets, the scores used are actually deviations from a target's mean. Hence the DA component actually confounds two types of accuracy: (1) the accuracy with which a judge predicts the score of an item for a particular target relative to the central tendency of other item scores for that target; and (2) the accuracy with which the judge predicts which target has the highest of these relative scores for each item.

It seems reasonable to argue that fidelity in assessing accuracy for a single target should be demonstrated before attempting to assess judges' abilities to differentiate among several targets. Also, it would appear that this second type of accuracy would to some extent depend on the first. Thus if one were interested in both types of accuracy, a sequential strategy would seem more appropriate. Finally, from another point of view, these two types of accuracy might be considered very distinct, and actually concerned with two entirely independent types of interpersonal judgment. Regardless of one's point of view, it seems clear that the two types of abilities certainly should not be confounded in one score. The model for inferential accuracy (Jackson, 1972; Reed & Jackson, 1975) eliminates this problem by considering only accuracy in judging the relative probabilities of behaviours for a single target. This type of accuracy has been labeled sensitivity. Because it deals with a single type of judgment and accuracy, it is believed to represent a much more psychologically meaningful process.

A second source of difficulty with Cronbach's components is the meaning of stereotype accuracy (SA). Cronbach (1955, p. 179) states that SA

> "....might be called accuracy in predicting the generalized other. This score depends on J's knowledge of the relative frequency or popularity of the possible responses."

The expression in Figure 4 for SA indicates that the "frequency or popularity of the possible responses" referred to by Cronbach is based solely on the targets in a particular study. This means that stereotype accuracy reflects the accuracy in predicting the generalized other or average person only to the extent that the average person is estimated by the sample of targets in a particular study. Clearly, sensitivity to the mean of the sample of targets may be quite different from sensitivity to the mean of people in general. This potential difference between the population mean and the mean of the targets in a study creates further difficulties in the interpretation of the components of Cronbach's analysis.

Cronbach advocated the removal of stereotype accuracy from the general accuracy score because its presence can artificially inflate accuracy correlations. In general principle, I certainly support the concept of unconfounding distinct components and removing biases from accuracy correlations. However, removing from an accuracy score what Cronbach calls stereotype accuracy could in some instances remove the most meaningful or important component of accuracy. This would occur, for example, when the set of targets is quite homogeneous in some respect, such as when all targets are psychiatric patients. In this case, the mean of the targets would represent a general factor of psychopathology, and thus would represent an extremely important and meaningful component of the interpersonal judgment task. In this instance one might want to consider this component separately, but one would not want to arbitrarily remove it.

From the above example one can see that the removal of SA from the accuracy score may not produce the desired results. Cronbach was profound in pointing out the necessity of making distinct the concepts of sensitivity to the population mean and sensitivity to the behaviour of a single individual. The problem is that, in most cases, his formula for SA will not accomplish this. By using behaviour exemplars for which the population response frequencies are known, the model for inferential accuracy is easily able to accomplish what the removal of SA was intended to do. It is a simple matter to compute a profile of behavioural probabilities representing the population means for behavioural dimensions. This population profile can then be used to statistically remove from a judge's sensitivity score the effects of high sensitivity to the population frequencies of behaviours. In order to see more easily how this is accomplished it is necessary to look more closely at the model for inferential accuracy.

As mentioned above, the Cronbach (1955) analysis of components of accuracy scores utilized a squared difference or distance approach to comparing the predicted and criterion responses. Basically this was identical to the Cronbach and Gleser (1953) distance approach to the assessment of similarity between two profiles. In the latter paper Cronbach (1955) simply extended the model to include multiple sets of pairs of profiles (i.e., the number of targets was allowed to be greater than 1). In both cases the aim was the analysis of the global squared distance (D^2) score into components expressed as functions of elevation, scatter, and shape. Elevation being the mean level of response across variables or items, scatter referring to the variability of scores about the mean, and shape indicating the pattern of ups and downs across variables in the profile.

A number of investigators (Lorr, 1966; Overall & Klett, 1972) have suggested a vector-product approach as an alternative to the distance model for assessing profile similarity. One such approach is to use the correlation (r) as an index of similarity. An argument in favour of this practice is that D^2 confounds elevation, scatter, and shape, thus making a univocal interpretation of a given value of D^2 impossible. Obviously, however, there are great disadvantages associated with r as well, namely the discarding of the information contained in the elevation and scatter parameters. Skinner (1975; in press) has proposed a sequential vector-product strategy which isolates the individual contributions of elevation, scatter, and shape to the overall similarity between two profiles. This overall similarity is expressed as the average raw cross-products, CP, between the elements of the two profiles:

$$CP_{xy} = \frac{\sum_{i=1}^{n} x_i y_i}{n}$$

Following Horn (1969), Skinner (1975, 1977) shows that this term can be expressed as

$$CP_{xy} = m_x m_y + s_x s_y r_{xy}$$

where m, s, and r stand for mean, standard deviation, and correlation respectively. Given this equation, it is a trivial matter to compute the individual contributions of elevation, scatter and shape to the overall similarity, CP, between the two profiles.

The implications this has for the model for inferential accuracy are obvious. Recall that the revised model has three parameters: elevation, dispersion, and sensitivity. Sensitivity is the similarity in shape between the profile of judgment and the criterion profile. Dispersion is simply the scatter parameter for the judgment profile. Elevation is the profile mean. Using CP as the index of similarity between the criterion and judgment profiles one can completely account for the overall similarity from a knowledge of the values of the elevation and dispersion parameters for the two profiles along with the judge's sensitivity, which is simply the correlation between the two profiles.

Summary

Comparing the model for inferential accuracy using a vector-product method of assessing overall accuracy with Cronbach's D^2 approach one finds the following:

(1) Using the vector-product index CP means that a large value for CP indicates a high level of accuracy, while using D^2 means that a high value for D^2 indicates a low degree of accuracy.

(2) The formulae for the components of accuracy using the vector-product model are much simpler and direct than those of the D^2 approach.

(3) The elimination of consideration of multiple targets in conjunction with the more direct vector-product formulae makes the sensitivity parameter of the model for inferential accuracy much more easily derived and understood than its counterpart, the DA correlation term.

(4) More important than 2 and 3 is the fact that the components of overall accuracy correspond to meaningful psychological processes which can be subjected to experimental investigation. In addition, the components do not confound different types of judgment as is the case with DA.

(5) Finally, it is a simple matter to remove the effects of sensitivity to the average person or population base rates for behaviours (what Cronbach meant by "stereotype accuracy"). First, one partials the profile of base rates from the judgment and criterion profiles. Conceptually this involves obtaining for each of these two profiles a linear transformation of the original profile which is the best least

squares approximation to the original profile with the restriction that it have a correlation of zero with the base rate profile. The correlation of these new profiles will reflect the judge's sensitivity independent of stereotype accuracy. Most important here is the fact that this procedure actually accomplishes what the D^2 formulas intended but failed to accomplish.

The analytical procedures having been developed, it is now appropriate to return to the criterion issue. The establishment of the validity of the criterion profiles used in the model for inferential accuracy should make the model a powerful tool for investigating the inferential process. It was with this in mind that the empirical investigations reported in the next chapter were undertaken.

CHAPTER EIGHT

EMPIRICAL INVESTIGATIONS

A series of investigations has been conducted looking at the reliability, accuracy, generalizability of accuracy, and components of the group consensus judgment. In addition, a study is reported concerning the generalizability of individual sensitivity scores across both targets and tests. Each of these studies has incorporated the model modifications just described.

The basic approach for each major investigation involved a cross-validation or replication approach. That is, each investigation actually consisted of two simultaneous but independent studies involving different subjects and different items. For the remainder of this report the convention will be adopted of referring to the major investigations as such (e.g., Investigation II), and referring to the simultaneous studies within each investigation as Studies A and B.

Each investigation was based on a large data pool collected by the author. The data collection was designed to accommodate each of the reported investigations. The subjects, the selection and development of targets, choice of items, and the design and method of data collection are described next. Because each investigation is based on the same data pool, to avoid repetition this information is being presented once, as an introduction to the individual investigations.

Subjects

The subjects used in these investigations were 208 (110 females, 98 males) introductory psychology students receiving required research credits for participation. Four subjects were discarded because of missing responses (more than 1% missing) yielding a final n of 204.

Targets

Modal Profile Analysis was completed for samples of PRF and DPI data. These analyses have been described above, and the resulting modal profiles were presented in Tables 1b and 2b. Two modal profiles were chosen for each test, and descriptions were written for each highlighting behaviours representative of the salient behavioural dimensions for the respective modal profiles.

In the case of the PRF, target descriptions were based on PRF Modal Profiles I+ and II-. The DPI target descriptions were based on DPI Modal Profiles II+ and III-. The salient scales for each are presented in Table 6. The complete target descriptions are found in Appendix A.

Items

To produce meaningful and reliable scale scores, it is necessary to use considerably more items per scale than were used by Reed and Jackson (1975). The ideal, of course, would be to have each subject judge the entire test for each target. However, the number of judgments called for would require far too much time. As a compromise, it was decided that eight items per scale, four true-keyed and four false-keyed, would be used.

The design requirement of replicating studies calls for two comparable sets of items for each scale of both tests. Therefore, each 16-item PRF-E and DPI scale was divided into two 8-item scales (4 true-keyed and 4 false-keyed) which were matched as well as possible for frequency of endorsement, or p-value, in the keyed direction. Thus, each test was divided into two shorter tests which might be considered

Table 6

Salient Modal Profile Scales for Target Descriptions

PRF Targets	
Richard Hall (I+)	Ed Murdock (II-)
Understanding	Autonomy
Achievement	Aggression
Endurance	Dominance
Nurturance	Defendence
DPI Targets	
Jack Cole (II+)	Jim Anderson (III-)
Rebelliousness	Desocialization
Shallow Affect	Repression
Socially Deviant Attitudes	Self Depreciation
Sadism	Depression

parallel forms. These shorter forms are called APRF, BPRF, ADPI, and BDPI.

Data Collection

Subjects were run in several groups ranging in size from 8 to 70. For each group, two sessions were required taking place on consecutive days. On the first day, half of each group judged DPI targets and half judged PRF targets. On the second day, the targets were switched. For example, those subjects who judged DPI targets on day one, judged PRF targets on the second day. The order of a pair of targets was randomly determined for each subject.

In addition, the subjects were divided randomly into two approximately equal groups which are labeled Group A (n = 99) and B (n = 105). These labels are consistent with the item sets (A or B) which the group members judged. Both groups judged the same targets, but used different sets of items (see Figure 5). That is, Group A judged the likelihood that Jack Cole and Jim Anderson would endorse the ADPI items, while Group B judged the likelihood that Cole and Anderson would endorse the BDPI items. Group B also judged the likelihood that Richard Hall and Ed Murdock would endorse the BPRF items, while Group A made the same judgment for the APRF items. Remember also, that the test and target orders were randomly determined for each subject.

Criteria

Two criteria were examined as bases for assessing the accuracy of the group consensus (GC). To simplify the explanation of the criteria, each will be described in the context of Richard Hall, one of the PRF targets.

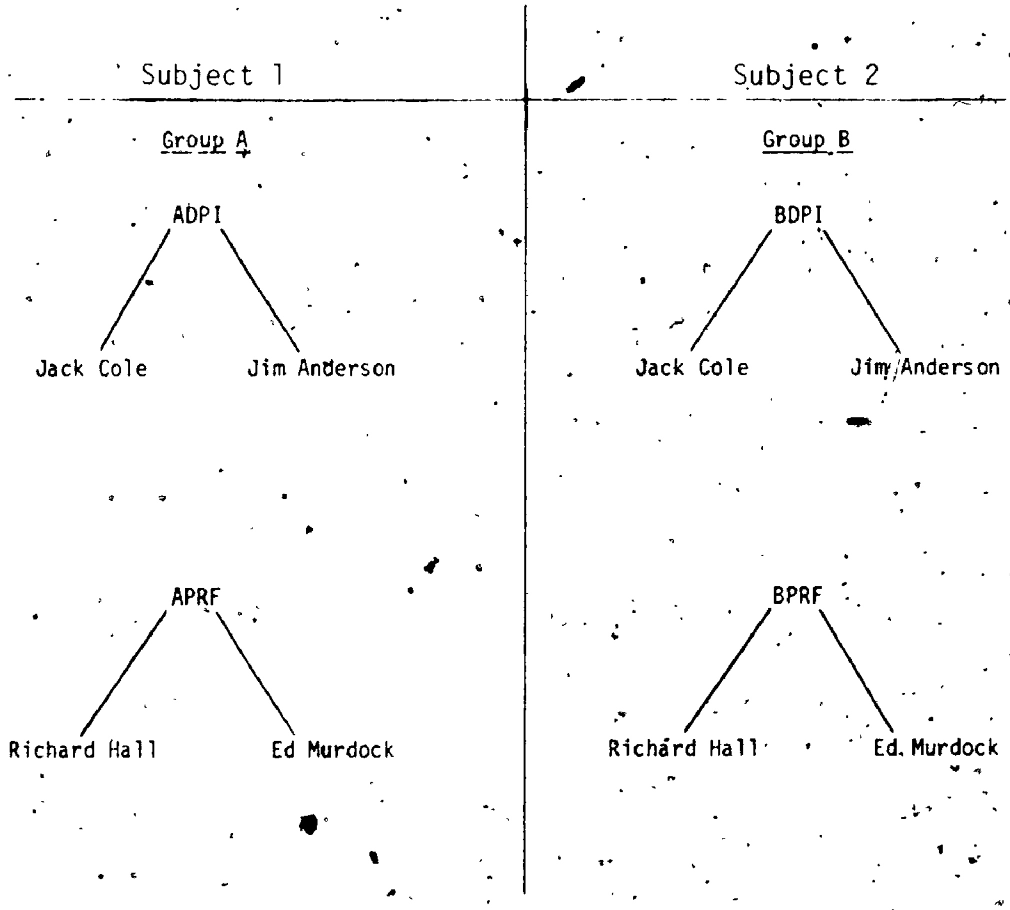


Figure 5. Design of Data Collection: Possible Orders for Two Hypothetical Subjects.

The first criterion is essentially the mean of the raw responses of 7 individuals (males) whose profiles are similar to Richard Hall's profile, that is, PRF-E Modal Profile I+. Similarity is defined as a correlation exceeding .50 (each scale score is first transformed to a standard score). The identified individuals are called target exemplars. The average frequency of endorsement for each PRF-E item was computed for this group of Richard Hall exemplars. Next, the items were divided into the APRF and BPRF item sets. It was then possible to compute scale scores for each PRF scale, within each item set. This yielded two sets of criterion PRF scale scores for Richard Hall exemplars, one for the APRF and one for the BPRF. The final step was to compute the mean APRF and BPRF profiles for these Hall exemplars. It is the mean profiles which were used as the first objective criterion for the Richard Hall target.

This procedure was repeated for each of the targets. The PRF target exemplars were selected from a sample of male college students. The DPI target exemplars were selected from samples of prison inmates who had undergone psychological assessment. Each exemplar criterion profile used in subsequent analyses is actually the mean of 7 target exemplars computed just as described above for the case of Richard Hall exemplars.

To describe the second criterion we will once again use the example of Richard Hall. The second criterion is Richard Hall's modal profile itself (Modal Profile I+, Table 2b). These profile scores represent the relative behavioural probabilities for the individuals defining the positive pole of the first factor of the PRF-E Modal Profile Analysis, and are the scores that determined which behaviours would be emphasized in the description of Richard Hall. An advantage of the modal profile criterion is that different profiles are statistically independent. That is, the respective PRF modal profiles for the Ed Murdock and Richard Hall targets have a correlation equal to zero. The same is true for the modal profiles of DPI targets.

Reliability and Convergent and Discriminant Validity of the Criterion Profiles

Before using the criterion profiles to access the accuracy of the group consensus profile, it was first necessary to investigate the reliabilities and intercorrealtions of the various criterion profiles. These analyses for PRF and DPI criterion profiles are reported in Table 7. As one can see, combining the A and B studies produces three different profiles for each target. For example, for the Richard Hall target there is the APRF exemplar profile, the BPRF exemplar profile, and the modal profile on which the description of Richard Hall was based. Looking first at the top half of Table 7, that dealing with the PRF criterion profiles, one should note the extremely high correlation (.979) between the APRF and BPRF profiles for the Richard Hall exemplars. The underlined correlations can be interpreted as strong tests of the reliability of the exemplar profiles because they represent a reliability across short parallel forms. The correlation for Richard Hall reflects the virtually identical shape of the profiles based on the two parallel forms of the PRF. Looking at the other underlined correlations one finds that the reliabilities for the Murdock and Cole exemplar profiles are also very high (.903 and .893) and for Anderson it is somewhat lower (.747), but still quite satisfactory.

The reliability coefficient for each target in conjunction with two other intra-target correlations constitute a test of the convergent validity of the criterion profiles. These three correlations form

Table 7

Reliability and Convergent and Discriminant Validity of Criterion Profiles for PRF and DPI Studies[a,b]

	PRF				
	Hall Exemplars[b] APRF	Hall Exemplars BPRF	Hall Modal Profile	Murdock Exemplars APRF	Murdock Exemplars BPRF
Hall Exemplars BPRF	.979				
Hall Modal Profile	.915	.933			
Murdock Exemplars APRF	.554	.539	.406		
Murdock Exemplars BPRF	462	.438	.382	903	
Murdock Modal Profile	.095	.077	000	.769	.795

	DPI				
	Anderson Exemplars ADPI	Anderson Exemplars BDPI	Anderson Modal Profile	Cole Exemplars ADPI	Cole Exemplars BDPI
Anderson Exemplars BDPI	.747				
Anderson Modal Profile	.705	.713			
Cole Exemplars ADPI	.154	-.100	-.234		
Cole Exemplars BDPI	.226	.041	-.130	.893	
Cole Modal Profile	.180	.077	-.001	.833	.841

[a]Correlations reflect similarity of criterion profile shape across scales

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target.

Table 9

Generalizability and Discriminant Validity of Group Consensus Profiles for PRF and DPI Targets

PRF

	Richard Hall Group Consensus: APRF	Richard Hall Group Consensus: BPRF	Ed Murdock Group Consensus: APRF
Richard Hall Group Consensus: BPRF	.944		
Ed Murdock Group Consensus: APRF	-.274	-.162	
Ed Murdock Group Consensus: BPRF	-.512	-.421	.921

DPI

	Jim Anderson Group Consensus: ADPI	Jim Anderson Group Consensus: BDPI	Jack Cole Group Consensus: ADPI
Jim Anderson Group Consensus: BDPI	.946		
Jack Cole Group Consensus: ADPI	-.312	-.262	
Jack Cole Group Consensus: BDPI	-.441	-.399	.922

Note - A and B group consensus profiles based on 99 and 105 judges respectively.

discriminant validity, the lower the correlation the better. Of course the modal profiles are defined as uncorrelated and in fact were found to be so empirically. Ideally, the remaining inter-target correlations will also approach zero, thus confirming the discriminant validity of the criterion profiles.

Table 7 might suggest that the discriminant validity for the PRF criterion profiles is somewhat less than ideal. One notes in Table 8, however, the reduction in the intercorrelations of the PRF criterion profiles when the effects of the population means are removed. The fact that the inter-target correlations are higher for Table 7 than is the case in Table 8 simply reflects the fact that differences among the population means for the PRF scales will artifactually inflate the correlations between profiles. Thus, Table 8 shows that with the removal of the effects of the population means, the PRF criterion profiles are discriminantly valid.

The discriminant validity of the DPI criterion profiles is evident even in Table 7 as indicated by the very small inter-target correlations. Obviously this also indicates that differences among population means for DPI scales had very little effect on the correlations among criterion profiles. This is consistent with the rather minimal effects of partialing out the population means as can be seen upon inspection of the lower half of Table 8.

Tables 7 and 8 clearly demonstrate the reliability of the criterion profiles as well as their convergent and discriminant validity. These criterion properties are important in that they strengthen the basis for conclusions which are drawn from each of the four investigations reported below.

Table 8

Reliability and Convergent and Discriminant Validity of Criterion Profiles: Population Means Partialed Out[a,b]

PRF

	Hall Exemplars[b] APRF	Hall Exemplars BPRF	Hall Modal Profile	Murdock Exemplars APRF	Murdock Exemplars BPRF
Hall Exemplars BPRF	845 (98)				
Hall Modal Profile	772 (91)	790 (93)			
Murdock Exemplars APRF	293 (59)	293 (54)	06[illegible] (.41)		
Murdock Exemplars BPRF	242 (.45)	205 (44)	120 (38)	897 (90)	
Murdock Modal Profile	256 (10)	249 (.08)	068 (00)	912 (77)	928 (80)

DPI

	Anderson Exemplars ADPI	Anderson Exemplars BDPI	Anderson Modal Profile	Cole Exemplars ADPI	Cole Exemplars BDPI
Anderson Exemplars BDPI	761 (.75)				
Anderson Modal Profile	.82[illegible] (.70)	.757 (71)			
Cole Exemplars ADPI	062 (15)	-.112 (-.10)	-.033 (-.23)		
Cole Exemplars BDPI	.177 (.23)	051 (.04)	.052 (-.13)	850 (.89)	
Cole Modal Profile	.139 (.18)	.086 (08)	.121 (00)	.854 (83)	.825 (.84)

[a]Correlations reflect similarity of criterion profile shape across scales, with the influence of the population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target.

Investigation I.

Accuracy and Generalizability of the Group Consensus Judgment

The Reed and Jackson (1975) study has already demonstrated the reliability of the group consensus judgment across sets of judges. The design of the present study, however, makes possible the testing of a different type of reliability, perhaps more appropriately called generalizability. In the present investigation it is possible to assess the generalizability of the group consensus judgment across both subjects and items. Group A and Group B recall, differ not only with respect to the individuals constituting each group, but also with respect to the items that were judged for each test (i.e., APRF and BPRF respectively). Therefore, if one were to compute the correlation between the Richard Hall group consensus profile for the APRF group and the Richard Hall group consensus profile for the BPRF group, one could obtain a coefficient of generalizability of the group consensus judgment across both individual judges and items being judged for a given target. These coefficients were computed for both of the PRF targets (Hall and Murdock) as well as for the DPI targets (Anderson & Cole), and are found in the upper left and lower right hand portions of the appropriate sections of Table 9 (the underlined values). The fact that the lowest coefficient is .921 (Murdock) is even more striking than the .99 reliability reported by Reed and Jackson (1975) when one considers that the present coefficients reflect the generalizability across both judges and items. These high coefficients of generalizability indicate not only that the judgments are dependable, but also attest to the utility of the rationale and strategy underlying the construction of the PRF and the DPI.

Table 9

Generalizability and Discriminant Validity of Group
Consensus Profiles for PRF and DPI Targets

PRF

	Richard Hall Group Consensus: APRF	Richard Hall Group Consensus: BPRF	Ed Murdock Group Consensus: APRF
Richard Hall Group Consensus: BPRF	.944		
Ed Murdock Group Consensus: APRF	-.274	-.162	
Ed Murdock Group Consensus: BPRF	-.512	-.421	.921

DPI

	Jim Anderson Group Consensus: ADPI	Jim Anderson Group Consensus: BDPI	Jack Cole Group Consensus: ADPI
Jim Anderson Group Consensus: BDPI	.946		
Jack Cole Group Consensus: ADPI	-.312	-.262	
Jack Cole Group Consensus: BDPI	-.441	-.399	.922

Note - A and B group consensus profiles based on 99 and 105 judges respectively.

Also reported in Table 9 are the inter-target group consensus correlations. For example, the correlation between the Richard Hall APRF group consensus profile and the Ed Murdock group consensus profile is -.274. As was the case for the criterion profiles, these inter-target correlations are an indication of the discriminant validity or independence of the profiles, although in this instance it is the judgment profiles we are dealing with. In the cases of both the PRF and the DPI these correlations are moderate and negative, with means of -.342 and -.354 respectively. Compared with convergent validities in excess of .92, these correlations substantiate the discriminant validity of the group consensus profiles.

The magnitude of the PRF correlations reduces when the effects of the population means is removed (see Table 10). Recall this is consistent with the comparable criterion correlations. In Table 11 one observes that with desirability partialed out the inter-target correlations for the PRF group consensus profiles are reduced even more substantially (discriminant validity increases) while in the case of the DPI inter-target group consensus correlations, the magnitude actually increases slightly. The reasons for these differences include the possibility that judges have more poorly differentiated networks for psychopathological behaviours and will be discussed further in the next section.

Also note in Tables 10 and 11 that the removal of the population means and the desirability variance does little to effect the generalizability coefficients, with the exception of the Richard Hall target. This generalizability coefficient drops from its original zero-order correlation of .944 to .824 with the population means partialed out,

Table 10

Generalizability and Discriminant Validity of Group Consensus

Profiles for PRF and DPI Targets:

Population Means Partialed Out

PRF

	Richard Hall Group Consensus: APRF	Richard Hall Group Consensus: BPRF	Ed Murdock Group Consensus: APRF
Richard Hall Group Consensus: BPRF	.824 (.94)[a]		
Ed Murdock Group Consensus: APRF	-.206 (-.27)	-.031 (-.16)	
Ed Murdock Group Consensus: BPRF	-.297 (-.51)	-.176 (-.42)	.940 (.92)

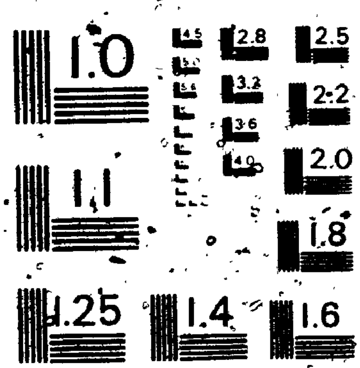
DPI

	Jim Anderson Group Consensus: ADPI	Jim Anderson Group Consensus: BDPI	Jack Cole Group Consensus: ADPI
Jim Anderson Group Consensus: BDPI	.942 (.95)[a]		
Jack Cole Group Consensus: ADPI	-.240 (-.31)	-.190 (-.26)	
Jack Cole Group Consensus: BDPI	-.387 (-.44)	-.360 (-.40)	.887 (.92)

Note - A and B group consensus profiles based on 99 and 105 judges respectively.

[a]Values in parentheses are the original unpartialed correlations.

2 OF/DE 3



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963

ARC 374

Table 11

Generalizability and Discriminant Validity of Group Consensus

Profiles for PRF and DPI Targets:

Desirability Partialed Out

PRF

	Richard Hall Group Consensus: APRF	Richard Hall Group Consensus: BPRF	Ed Murdock Group Consensus: APRF
Richard Hall Group Consensus: BPRF	.734 (.94)[a]		
Ed Murdock Group Consensus: APRF	-.124 (-.27)	-.083 (-.16)	
Ed Murdock Group Consensus: BPRF	-.189 (-.51)	-.011 (-.42)	.952 (.92)

DPI

	Jim Anderson Group Consensus: ADPI	Jim Anderson Group Consensus: BDPI	Jack Cole Group Consensus: ADPI
Jim Anderson Group Consensus: BDPI	.929 (.95)[a]		
Jack Cole Group Consensus: ADPI	-.463 (-.31)	-.348 (-.26)	
Jack Cole Group Consensus: BDPI	-.565 (-.44)	-.476 (-.40)	.923 (.92)

Note - A and B group consensus profiles based on 99 and 105 judges respectively.

[a]Values in parentheses are the original unpartialed correlations.

Table 13a

PRF Group Consensus Accuracy Correlations for Two Targets:

Population Means Partialed Out[a]

Criterion Profiles	APRF Study n = 99: Richard Hall Group Consensus Profile	APRF Study n = 99: Ed Murdock Group Consensus Profile	BPRF Study n = 105: Richard Hall Group Consensus Profile	BPRF Study n = 105: Ed Murdock Group Consensus Profile
Hall Exemplars[b]	.250 (.70)	.165 (-.07)	.492 (.76)	.005 (-.35)
Hall Modal Profile	<u>.464</u> (.77)	-.100 (-.20)	<u>.609</u> (.85)	-.210 (-.45)
Murdock Exemplars[b]	-.362 (.13)	.818 (.61)	.063 (.28)	.751 (.47)
Murdock Modal Profile	-.535 (-.41)	<u>.769</u> (.76)	-.226 (-.19)	<u>.799</u> (.75)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 21 PRF scales, with the influence of population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 13b

DPI Group Consensus Accuracy Correlations for Two Targets:

Population Means Partialed Out[a]

Criterion Profiles	ADPI Study n = 99: Jim Anderson Group Consensus Profile	ADPI Study n = 99: Jack Cole Group Consensus Profile	BDPI Study n = 105: Jim Anderson Group Consensus Profile	BDPI Study n = 105: Jack Cole Group Consensus Profile
Anderson Exemplars[b]	.511 (.45)	-.209 (-.50)	<u>.449</u> (.44)	-.321 (-.30)
Anderson Modal Prof	<u>.571</u> (.60)	-.253 (-.54)	<u>.492</u> (.52)	-.182 (-.30)
Cole Exemplars[b]	-.469 (-.10)	.727 (.79)	-.497 (-.52)	.726 (.78)
Cole Modal Profile	-.602 (-.37)	<u>.532</u> (.59)	-.485 (-.51)	<u>.669</u> (.71)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 25 DPI scales, with the influence of population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 12a

PRF Group Consensus Accuracy Correlations for Two Targets:

Exemplar and Modal Profile Criteria[a]

Criterion Profiles	APRF Study n = 99		BPRF Study n = 105	
	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile
Hall Exemplars[b]	.695	-.073	.76[illegible]	-.347
Hall Modal Profile	.769	-.204	.849	-.448
Murdock Exemplars[b]	.127	.607	.278	.469
Murdock Modal Profile	-.411	.763	-.188	.754

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 21 PRF scales.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 12b

DPI Group Consensus Accuracy Correlations for Two Targets:

Exemplar and Modal Profile Criteria[a]

Criterion Profiles	ADPI Study n = 99		BDPI Study n = 105	
	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile
Anderson Exemplars[b]	.450	-.502	.443	-.298
Anderson Modal Prof	.5[illegible]9	-.535	.523	-.303
Cole Exemplars[b]	-.105	.790	-.519	.775
Cole Modal Profile	-.271	.593	-.514	.710

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 25 DPI Scales.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

the effect of desirability, and removing the effect of both simultaneously. These analyses are reported in Tables 13 and 14, and Appendix B, respectively. It should be kept in mind that the APRF and ADPI studies utilize the same 99 subjects, and that the BPRF and BDPI studies utilize a different group of 105 subjects.

The correlations reported in Tables 12 to 14 reflect the similarity in shape of the group consensus profile with respect to each of the criterion profiles (Exemplars and Modal Profile). In those cases where partial correlations are reported, the correlations reflect the similarity in shape between a given group consensus profile and a criterion profile with the effects of either the population means for the scales or desirability removed. The accuracy correlations are found in the main diagonal of the values recorded for each study. These values are underlined in each table. The off-diagonal entries for each study represent the correlation between a target group consensus profile and the criterion profile for a different target. These correlations are an indication of the discriminant validity of the accuracy scores.

Looking at the APRF and BPRF studies (Table 12a) one finds substantial accuracy for both the Richard Hall group consensus profile and the Ed Murdock group consensus profile as measured with respect to each of the criterion profiles. The similarity of the results of the APRF and BPRF studies is evident. The results of this APRF study with its BPRF replication offer strong support to the assumption in the model that the group consensus represents the actual covariation of behaviours in the target individuals. The DPI studies (Table 12b) offer corroborating findings. The accuracy correlations for the Jack Cole target are similar in magnitude to those accuracy correlations for both PRF targets. The accuracy of the Jim Anderson group consensus profiles is somewhat lower, but considering the expected increase in the difficulty of judging psychopathological behaviours, and expected attenuation due to the lower reliabilities of the criterion and group consensus profiles for the DPI targets, these accuracy correlations provide strong support for the idea that there is considerable accuracy in the group consensus judgment.

Tables 13 and 14 report the group consensus accuracy correlations with the effects of the population means and the effects of desirability removed respectively. The values in parentheses are the original zero-order correlations reported in Table 12. In these tables one finds a dramatic effect on the Richard Hall correlations with the removal of either the variance associated with the means or the variance associated with desirability. The Richard Hall group consensus accuracies dropped considerably with the removal of the population means (Table 13a) and to an even greater extent with the removal of the desirability component of the variance (Table 14a). However, the effects on the remaining accuracy correlations were quite different, particularly in the cases in which desirability was partialed out (Tables 14a & b).

Consider first the Ed Murdock accuracy correlations in Table 13a. One notes an increase in the correlation of the group consensus with the Murdock exemplar profile from .61 to .818 for the APRF Study and from .47 to .75 for the BPRF Study. Again this reflects the differences

Table 13a

PRF Group Consensus Accuracy Correlations for Two Targets:

Population Means Partialed Out[a]

	APRF Study n = 99		BPRF Study n = 105	
Criterion Profiles	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile
Hall Exemplars[b]	.250 (.70)	.165 (-.07)	.492 (.76)	.005 (-.35)
Hall Modal Profile	.464 (.77)	-.100 (-.20)	.609 (.85)	-.210 (-.45)
Murdock Exemplars[b]	-.362 (.13)	.818 (.61)	.063 (.28)	.751 (.47)
Murdock Modal Profile	-.535 (-.41)	.769 (.76)	-.226 (-.19)	.799 (.75)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 21 PRF scales, with the influence of population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 13b

DPI Group Consensus Accuracy Correlations for Two Targets:

Population Means Partialed Out[a]

	ADPI Study n = 99		BDPI Study n = 105	
Criterion Profiles	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile
Anderson Exemplars[b]	.511 (.45)	-.209 (-.50)	.449 (.44)	-.321 (-.30)
Anderson Modal Prof	.571 (.60)	-.253 (-.54)	.492 (.52)	-.182 (-.30)
Cole Exemplars[b]	-.469 (-.10)	.727 (.79)	-.497 (-.52)	.726 (.78)
Cole Modal Profile	-.602 (-.37)	.532 (.59)	-.485 (-.51)	.669 (.71)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 25 DPI scales, with the influence of population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 14a

PRF Group Consensus Accuracy Correlations for Two Targets.
Desirability Partialed Out[a]

	APRF Study n = 99		BPRF Study n = 105	
Criterion Profiles	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile
Hall Exemplars[b]	.052 (.70)	.225 (-.07)	.389 (.76)	.156 (-.35)
Hall Modal Profile	[illegible] (.77)	-.011 (-.20)	.409 (.85)	-.055 (-.45)
Murdock Exemplars[b]	-.252 (.13)	.742 (.61)	.081 (.28)	.725 (.47)
Murdock Modal Profile	-.463 (-.41)	.751 (.76)	-.098 (-.19)	.787 (.75)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 21 PRF scales, with the influence of population means removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Table 14b.

DPI Group Consensus Accuracy Correlations for Two Targets
Desirability Partialed Out[a]

	ADPI Study n = 99		BDPI Study n = 105	
Criterion Profiles	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile
Anderson Exemplars[b]	.585 (.45)	-.086 (-.50)	.629 (.44)	[illegible] (-.30)
Anderson Modal Prof.	[illegible] (.60)	-.421 (-.54)	.489 (.52)	-.329 (-.30)
Cole Exemplars[b]	-.429 (-.10)	.894 (.79)	-.417 (-.52)	.861 (.78)
Cole Modal Profile	-.517 (-.37)	[illegible] (.59)	-.443 (-.51)	.754 (.71)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 25 DPI scales, with the influence of desirability removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

between the college sample population means from which the Ed Murdock exemplars were drawn and the general population means for the PRF scales. Note in particular that the Ed Murdock group consensus accuracy correlations hold up exceptionally well when the population mean and desirability variance is removed (Tables 13a & 14a). These correlations, the lowest being .725, reflect the very high level of accuracy which is present in the judgments of the Ed Murdock target. This accuracy indicates a sensitivity to the content or substance of the behavioural exemplars (APRF & BPRF items) because it is independent of the effects of the base rates and the desirabilities of behaviours. This is a good illustration of the strategy described in the previous chapter for systematically removing unwanted sources of variance from the accuracy index.

Findings for the judgments of the DPI targets similarly reveal a high level of accuracy in the group consensus judgment, but as was expected these accuracy correlations were slightly lower than was the case for the PRF targets, suggesting again that judgments of psychopathological behaviours are more difficult than judgments of behaviours in the normal range. Again, we observe a slight increase in accuracy associated with the removal of the desirability variance (Table 14), but in contrast to the Ed Murdock studies, no such effect for population means. With respect to the off-diagonal elements for the DPI portions of Tables 13 and 14, one sees somewhat of an increase in the magnitude of these negative correlations over those reported in Table 12. These are to some extent puzzling, but suggest to the author that perhaps in the case of the DPI judgments, the subjects have a much less clearly differentiated inferential network. This would mean that subjects would not be able to differentiate types as distinctly as necessary and could easily lead to correlated judged profiles when the corresponding criterion profiles were independent. This reduced differentiation is probably due to subjects' relatively low level of experience with psychopathological behaviours as compared with their high level of experience with the behaviours measured by the PRF scales. This difference in experience with the domains of behaviours measured by the respective tests is also the most likely cause of the differences between the DPI and PRF Stud[illegible]n the level of accuracy achieved by the judges, and also the slight increase in the DPI inter-target group consensus partial correlations mentioned in previous sections. In conclusion, the examination of the group consensus accuracy correlations reported in Tables 12, 13, and 14 very strongly supports the conclusion that the group consensus judgment does reflect accurately differences regarding the pattern of behaviour probabilities for targets. Furthermore, the results for the Richard Hall target demonstrate that in some instances removal of base rate or desirability variance result in the loss of important and meaningful variance. As was suggested in Chapter Seven, one must carefully consider what one is dealing with before discarding any component of accuracy.

Table 16

Generalizability of Group Consensus Accuracy for PRF and DPI Targets: Population Means Partialed out

	Generalizability Accuracy r[a] (i.e., APRF Group Consensus Profile: BPRF Exemplar Profile)	Original Accuracy r[b] (i.e., APRF Group Consensus Profile: APRF Exemplar Profile)
Richard Hall APRF Group Consensus	.215 (.73)[c]	.250 (.70)
Richard Hall BPRF Group Consensus	.218 (.76)	.492 (.83)
Ed Murdock APRF Group Consensus	.751 (.60)	.818 (.61)
Ed Murdock BPRF Group Consensus	.779 (.42)	.755 (.47)
Jim Anderson ADPI Group Consensus	.433 (.42)	.511 (.45)
Jim Anderson BDPI Group Consensus	.415 (.37)	.449 (.44)
Jack Cole ADPI Group Consensus	.562 (.67)	.727 (.79)
Jack Cole BDPI Group Consensus	.788 (.83)	.726 (.78)

[a]Generalizability Accuracy r refers to the correlation of the target group consensus profile from one set of items (i.e., ADPI) with the exemplar profile for the other set of items (BDPI), with the influence of the population means removed.

[b]Original Accuracy r refers to the correlation of the target group consensus and exemplar profiles from the same item set (i.e., ADPI: ADPI), with the influence of the population means removed.

[c]Values in parentheses are the unpartialed (zero-order) correlations.

Table 15

Generalizability of Group Consensus Accuracy

for PRF and DPI Targets

	Generalizability Accuracy r[a] (i.e., APRF Group Consensus Profile: BPRF Exemplar Profile)	Original Accuracy r[b] (i.e., APRF Group Consensus Profile APRF Exemplar Profile)
Richard Hall APRF Group Consensus	.730	.695
Richard Hall BPRF Group Consensus	.761	.829
Ed Murdock APRF Group Consensus	.604	.607
Ed Murdock BPRF Group Consensus	.418	.469
Jim Anderson ADPI Group Consensus	.425	.450
Jim Anderson BDPI Group Consensus	.371	.443
Jack Cole ADPI Group Consensus	.670	.790
Jack Cole BDPI Group Consensus	.832	.775

[a]Generalizability Accuracy r refers to the correlation of the target group consensus profile from one set of items (i.e., ADPI) with the exemplar profile for the other set of items (BDPI).

[b]Original Accuracy r refers to the correlation of the target group consensus and exemplar profiles from the same item set (i.e., ADPI: ADPI).

will be referred to as 'common item correlations'. By comparing the two correlations in each row, one can easily see the degree to which the accuracy generalizes to an independently derived criterion profile. For example, in the first row one sees that the APRF group consensus for Richard Hall correlates .730 with the BPRF exemplar profile for Richard Hall. Looking at the second column in row 1, we see that the group consensus-exemplar common item correlation is .695. That is, for the APRF group consensus for the Richard Hall target, the generalizability correlation is actually higher than the common item correlation. In the second row of Table 15 we see the generalizability in original accuracy correlations for the BPRF group consensus profile for Richard Hall. In this case we find that the generalizability correlation is slightly lower than the original or common item correlation (.761 as compared with .829).

For each target, then, we have two tests of generalizability. First we have the A group consensus-B exemplar correlation (generalizability r) as compared with the A group consensus-A exemplar correlation (common item r), and second we have the B group consensus-A exemplar correlation (generalizability r) as compared with the B group consensus-B exemplar correlation (common item r). As one looks at these correlations for each target, the pattern becomes very clear. The generalizability of these accuracy correlations is impressive.

To get a better idea of the trend of these correlations, the author squared each of the correlations in column 1 and column 2. Then the difference between the squared correlations was computed for each row and the average of these differences was obtained. This value was .069 indicating that the difference in the amount of variance accounted for by the original as compared with the generalizability correlation is less than 7%. In addition, the computation of this average value included two sets of correlations for which the generalizability accuracy was actually greater than the accuracy of the original correlation (Richard Hall APRF group consensus and Jack Cole BDPI group consensus). If the differences between these two sets of correlations are considered to be 0, then the average difference of the squared correlations is equal to .051, meaning that the difference in the amount of variance accounted for by the original versus the generalizability correlation is approximately 5%. This is a very encouraging finding, indicating the highly stable nature of the judgments and the extreme fidelity with which both the criterion and judgments are being measured.

Table 16 reports the generalizability of the group consensus accuracy with the effects of the population means partialed out, and Table 17 reports the generalizability of the group consensus accuracy with desirability partialed out. As expected, the generalizability correlations for the Richard Hall group consensus profiles drop drastically, just as was the case with the original common item r's. However, the remainder of the generalizability correlations hold up very well, and in fact many actually increase. For example, with desirability removed (Table 17) all the generalizability accuracy r's increase substantially over their zero-order counterparts for all of the targets other than Richard Hall, as was also the case for the original accuracy correlations reported in Table 13 and reported again in the second column of Table 17. In the case of the partialing out of population means (Table 16) for Ed Murdock we see a substantial increase for both generalizability and original accuracy correlations, while for both DPI targets there

Table 16

Generalizability of Group Consensus Accuracy for PRF and DPI Targets: Population Means Partialed out

	Generalizability Accuracy r[a] (i.e., APRF Group Consensus Profile: BPRF Exemplar Profile)	Original Accuracy r[b] (i.e., APRF Group Consensus Profile: APRF Exemplar Profile)
Richard Hall APRF Group Consensus	.215 (.73)[c]	.250 (.70)
Richard Hall BPRF Group Consensus	.218 (.76)	.492 (.83)
Ed Murdock APRF Group Consensus	.751 (.60)	.818 (.61)
Ed Murdock BPRF Group Consensus	.779 (.42)	.755 (.47)
Jim Anderson ADPI Group Consensus	.433 (.42)	.511 (.45)
Jim Anderson BDPI Group Consensus	.415 (.37)	.449 (.44)
Jack Cole ADPI Group Consensus	.562 (.67)	.727 (.79)
Jack Cole BDPI Group Consensus	.788 (.83)	.726 (.78)

[a]Generalizability Accuracy r refers to the correlation of the target group consensus profile from one set of items (i.e., ADPI) with the exemplar profile for the other set of items (BDPI), with the influence of the population means removed.

[b]Original Accuracy r refers to the correlation of the target group consensus and exemplar profiles from the same item set (i.e., ADPI: ADPI), with the influence of the population means removed.

[c]Values in parentheses are the unpartialed (zero-order) correlations.

predicting the group consensus profile from the three predictor profiles. In Study A (Table 18) we see that the multiple correlation of the three predictor profiles with group consensus profiles for the APRF Richard Hall target is .84, indicating that the amount of variance accounted for (R^2) in this group consensus profile is approximately 71%.

The second category of information reported for Study A concerns four types of importance weights for the three predicted profiles in predicting the group consensus profile. The first type of importance weight is labelled Predictor r, and consists simply of the zero-order correlation of each predictor profile with the criterion profile (group consensus). The second type of importance weight reported is labelled Orthogonal Predictor r. To obtain these importance weights, the three predictor profiles were simultaneously transformed such that each one was the least squares estimate of the original profile, with the restriction that the three resulting profiles be uncorrelated with one another. The analytical procedure for obtaining these orthogonal predictor profiles involved the use of a principle components procedure outlined by Mulaik (1972). The advantage of having uncorrelated predictors is that the importance of each predictor can be directly interpreted. The most efficient and meaningful way of doing this is by looking at the squared orthogonal predictor correlation which is reported in the third column of each table (Orthogonal Predictor r^2). Each of these squared correlations represents the amount of variance accounted for by the respective predictor. Because the predictors are orthogonal, one can simply sum the squared correlations of each of the predictors with the criterion and obtain the total amount of variance is basically little or no change. The double partialed accuracy correlations are very similar to each of the first-order correlations. They are reported in Appendix B.

Implications of Investigations I and II

The results of Investigation I and Investigation II present a clear and consistent pattern supporting a position that there is considerable accuracy in the group consensus judgment. On the average, individuals do correctly infer the covariation of behaviours in a target individual as assessed with respect to two types of objective criteria. In addition, by measuring this accuracy at the level of scale scores it is possible to demonstrate that the accuracy of the consensus inference with respect to such behaviours is not limited to only one set of specific behaviour exemplars, but that it generalizes to an independent set of behaviour exemplars which are concerned with a common set of behaviour categories (scales). This finding provides supportive evidence regarding the legitimacy of three theoretically important concepts: (1) conceptualizing behaviour in terms of patterns of individual behavioural consistency; (2) using a model of measurement concerned with dimensions of behaviour; and (3) representing the patterns of consistency in terms of profiles of behaviour probabilities. It seems highly unlikely that relations as consistent, strong, and robust as those reported in Investigations I and II could occur if any of these concepts was incorrect.

Investigation III:
Relative Importance of Components of the Group Consensus Judgment

Now that Investigations I and II have demonstrated the validity of the group consensus, it is appropriate to examine the group consensus judgment more closely. Investigation III was designed to do this in two ways. First of all, the investigation sought to determine how well we understand the group consensus judgment. "How well" is defined as the amount of the group consensus profile variance which can be accounted for in a multivariate prediction paradigm (multiple correlation). The second aspect of Investigation III concerns the identification of the relative importance of three components of the group consensus judgment. The investigation involves eight separate analyses, two (Study A and Study B) for each of the four different targets.

Three different predictor profiles were selected for each of the four targets. These will each be described individually, using the case of the prediction of the Richard Hall APRF group consensus profile as an example. The first predictor profile is a profile of population means for the 21 APRF scales. The second predictor profile was a profile of APRF scale scores derived from the mean desirability rating for each APRF item. The final predictor profile is the mean of the APRF scale scores for the seven exemplars of the Richard Hall target. This is the same exemplar profile that was used as a criterion for the accuracy correlations of the Richard Hall APRF group consensus profile in Investigations I and II.

We know from the first two investigations that there is considerable common variance between the exemplar and group consensus profiles, In addition, we know from those studies reported as parts of Investigations I and II that there remains considerable common variance between the exemplar and group consensus profile even when the effects of the population means, or the effects of desirability have been removed (except of course in the cases of the Richard Hall group consensus profiles). It is recognized that there is some redundancy in now turning around and predicting the group consensus profile using the three predictors of population means, desirability and exemplar profile. However, it is thought that by using the different approach of considering the group consensus profile as a criterion, and looking at the components of variance of this criterion that are accounted for by each of the predictor profiles, it will be possible to answer directly one additional important question. This question is as follows: To what extent does each group consensus judgment represent a judgment reflecting the content or substantive element of behaviours, and to what extent does each group consensus judgment merely reflect a sensitivity to properties or components of behavioural exemplars which are extraneous to content, namely the base rate of behaviours (the population means profile) and the desirability of behaviours? In other words, Investigation III will provide us directly with information regarding the relative importance of three different components of the group consensus judgment for each target. These components are base rate, desirability, and content.

Tables 18 through 21 present the A and B studies for the group consensus profiles for Richard Hall, Ed Murdock, Jim Anderson, and Jack Cole targets respectively. Each table presents for its respective target, categories of information for Study A and Study B. First one finds the multiple correlation and the squared multiple correlation for predicting the group consensus profile from the three predictor profiles. In Study A (Table 18) we see that the multiple correlation of the three predictor profiles with group consensus profiles for the APRF Richard Hall target is .84, indicating that the amount of variance accounted for (R^2) in this group consensus profile is approximately 71%.

The second category of information reported for Study A concerns four types of importance weights for the three predicted profiles in predicting the group consensus profile. The first type of importance weight is labelled Predictor r, and consists simply of the zero-order correlation of each predictor profile with the criterion profile (group consensus). The second type of importance weight reported is labelled Orthogonal Predictor r. To obtain these importance weights, the three predictor profiles were simultaneously transformed such that each one was the least squares estimate of the original profile, with the restriction that the three resulting profiles be uncorrelated with one another. The analytical procedure for obtaining these orthogonal predictor profiles involved the use of a principle components procedure outlined by Mulaik (1972). The advantage of having uncorrelated predictors is that the importance of each predictor can be directly interpreted. The most efficient and meaningful way of doing this is by looking at the squared orthogonal predictor correlation which is reported in the third column of each table (Orthogonal Predictor r^2). Each of these squared correlations represents the amount of variance accounted for by the respective predictor. Because the predictors are orthogonal, one can simply sum the squared correlations of each of the predictors with the criterion and obtain the total amount of variance

Table 18

Multiple Correlation and Four Different Types of Importance Weights
in Predicting the Group Consensus Profile:
PRF Target Richard Hall[a]

Study A				
	R = .840		R^2 = .706	
	Importance Weights			
Predictor Profiles[b]	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	.708	.341	.117	-.031
Desirability	.840	.683	.466	.815
Exemplar	.695	.351	.123	.063

Study B				
	R = .871		R^2 = .759	
	Importance Weights			
Predictor Profiles[b]	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	.786	.393	.154	.002
Desirability	.846	.568	.323	.510
Exemplar	.829	.531	.282	.394

[a]It is the group consensus profile that is being predicted in the analyses presented in this table.

[b]Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

accounted for by the predictors. This can be verified by checking the sum of the weights in column 3 (Orthogonal Predictor r^2) with the value for R^2 reported in the upper right-hand corner for Study A.

The final type of importance weight is a traditional beta weight from the multiple correlation analysis. These are reported for the sake of completeness, as the author would argue following Darlington (1968) that there is considerable difficulty inherent in the attempt to interpret beta weights as importance weights. This author feels that the most directly interpretable weights in terms of importance weights are the Orthogonal Predictor r^2,s reported in the third column of the table. For the skeptical reader, the transformed factor loadings matrix from the orthogonalization procedure are reported in Appendix C. An examination of these loadings will show that the correlation between the original predictor profiles and the transformed predictor profiles are extremely high in all cases.

As would be expected, in the case of Richard Hall (Table 18) desirability accounts for the greatest portion of the variance of the group consensus profile. However, in Study A the exemplar profile does account for 12.3% of the variance indicating that a substantial amount of the group consensus judgment for the Richard Hall target can be given a univocal content interpretation. Study B produced an Orthogonal Predictor r^2 for the exemplar profile equal to .282, indicating that over twice as much variance is uniquely accounted for by the exemplar predictor than was the case in Study A. This shows that even in the case of the Richard Hall target, a target for which the accuracy of the group consensus fell greatly with the removal of the population mean and desirability variance, there remains a considerable component of content variance in the group consensus judgment.

The remainder of the Study B results reported in Table 18 are very consistent with the Study A results. The R^2 = .759 indicating that an even greater amount of the group consensus variance is accounted for in Study B than was the case in Study A.

Table 19 reports the multiple correlation and importance weights for the two Ed Murdock studies. Again one sees that the three components of means, desirability and exemplar account for a large proportion of the reliable group consensus variance. In this case the content (exemplar) component is clearly the largest of the three. By averaging the Orthogonal Predictors' r^2 from Study A and Study B for each of the components, one can get quite a reliable estimate of the relative importance of the respective components. The average content component is approximately 45% of the group consensus variance, while the average of the means and desirability components is 12 and 11% of the variance respectively. The average of the multiple R's is about .82.

The multiple correlations for the Jim Anderson DPI target (Table 20) are lowest of the four targets, but nonetheless are quite respectable (.707 & .751). In this case one finds that desirability and content components are of about equal importance (mean variance accounted for by Orthogonal Predictor r^2's over Studies A and B are 27.6% and 24.8% respectively) and that the population mean component is very small (mean Orthogonal Predictor r^2 less than .01).

The prediction of the Jack Cole group consensus profiles is the best of all four targets (Table 21). The multiple correlation for Study A is .925 and for Study B is .899. These are about as high

Table 19

Multiple Correlation and Four Different Types of Importance Weights in Predicting the Group Consensus Profile: PRF Target Ed Murdock[a]

Study A

$R = .825$ $R^2 = .681$

Predictor Profiles[b]	Importance Weights			
	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	-.186	-.323	.104	-.668
Desirability	-.249	-.196	.038	-.028
Exemplar	.607	.733	.538	.928

Study B

$R = .819$ $R^2 = .670$

Predictor Profiles[b]	Importance Weights			
	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	-.410	-.366	.134	-.388
Desirability	-.492	-.426	.182	-.351
Exemplar	.469	.595	.354	.721

[a] It is the group consensus profile that is being predicted in the analyses presented in this table.

[b] Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

Table 20

Multiple Correlation and Four Different Types of Importance Weights in Predicting the Group Consensus Profile: DPI Target Jim Anderson[a]

Study A

$R = .707$ $R^2 = .501$

Predictor Profiles[b]	Importance Weights: Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	-.230	-.127	.016	-.006
Desirability	-.490	-.498	.248	-.546
Exemplar	.450	.486	.236	.514

Study B

$R = .751$ $R^2 = .564$

Predictor Profiles	Importance Weights: Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	-.188	-.005	.000	-.253
Desirability	-.461	-.550	.303	-.743
Exemplar	.443	.511	.261	.609

[a]It is the group consensus profile that is being predicted in the analyses in this table.

[b]Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

Table 21

Multiple Correlation and Four Different Types of Importance Weights in Predicting the Group Consensus Profile.

DPI Target Jack Cole[a]

Study A

$R = .925$ $R^2 = .855$

Predictor Profiles[b]	Importance Weights			
	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	.453	.362	.131	.327
Desirability	-.174	-.381	.145	-.587
Exemplar	.790	.761	.579	.765

Study B

$R = .899$ $R^2 = .808$

Predictor Profiles[b]	Importance Weights			
	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	
Population Means	.397	.333	.111	.337
Desirability	-.048	-.295	.087	-.568
Exemplar	.775	.781	.611	.835

[a] It is the group consensus profile that is being predicted in the analyses presented in this table.

[b] Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

as the as the reliabilities of the profiles involved. An average of approximately 83% of the group consensus profile variance is being accounted for. Again in the case of Jack Cole, as was true also for the Ed Murdock group consensus profile, the content component is by far the largest. The average of the exemplar Orthogonal Predictor r^2 for Studies A and B is just under .60, while the means and desirability r^2's are each about .12.

Summary

Table 22 gives a final summary of the importance weights found in Investigation III. Averaged across the eight studies reported in Tables 18 through 21, the percent of pure content variance in the group consensus judgment is nearly four times as great as the variance attributable to base rates, and nearly twice as great as the variance attributable to desirability. When the two Richard Hall targets are not included the ratios of content to base rate and desirability are even higher.

The results of Investigations I & II indicate that high levels of accuracy are attributable to the group consensus judgment of each target. Investigation III provides information regarding the degree to which accuracy reflects a sensitivity to differences in the content or substance of the behaviour typically engaged in by individual representatives of each of the targets. In addition, this third investigation also provides an indication of the extent to which the group consensus judgment is understood. In this regard there are clearly differences among the targets. With respect to the Cole and Hall group consensus there are definitely no problems, as the average variance accounted for by the three components for these two targets is

Table 22

Percentage of Group Consensus Judgment Variance Accounted for by

Content, Desirability, and Base Rate

Component[a]	% of Group Consensus[b] Variance Accounted for: All Studies	% of Group Consensus[c] Variance Accounted for: Richard Hall Excluded
Content	37.3%	43.0%
Desirability	22.4%	16.7%
Base Rate	9.5%	8.3%

[a]Content is the average of the exemplar orthogonal r^2's, desirability the average of the desirability orthogonal r^2's, and base rate the average of the population means orthogonal r^2's.

[b]The average % of variance accounted for by each component for 8 studies (2 replications for each of 4 targets).

[c]The average % of variance accounted for by each component for 6 studies (Richard Hall studies excluded).

approximately 83% and 73% respectively. However, the average variance accounted for in the case of the Murdock group consensus profiles is slightly less (about 67%) and for Anderson it drops to 53%. While 53% is still high, the portion unaccounted for (47%) is without a doubt greater than for the other targets and is large enough to be a concern.

At least two possible explanations exist. Some of the decrease in predictability is probably caused by a larger error component due to the difficulty of the target. The salient dimensions for the Anderson target were much more affect oriented than was the case for any of the other targets. It seems reasonable to assume that inferences regarding affect would be more difficult than inferences about more overt behaviours. It is unlikely, however, that this could account for all of the difference. A second possible source of the reduced predictability for the Anderson group consensus might be the presence of more than one point of view about Anderson. The possibility of identifying distinct points of view for certain targets opens up an intriguing avenue for future research.

Investigation IV:

Individual Parameters and the Generalizability of Sensitivity Scores Across Targets and Tests

The accuracy of the average or consensual judgment having been demonstrated, and its components identified, the final step is to look at the distribution of the sensitivities of individual judges, and the generalizability of individual sensitivity across targets and tests. Appendix D contains the frequency distributions and statistics for the individual subject parameters (elevation, dispersion, sensitivity) for each of the four targets, as well as the generalizability correlations for elevation and dispersion. There are actually five parameters reported for each subject because sensitivity was computed three different ways, once with each of the three possible criteria: group consensus (GC), modal profile (MP), and exemplar (EX).

<u>Generalizability of Sensitivity Scores</u>

Table 23 reports the intercorrelations of the three sensitivity measures for each of the four targets. Again, two independent studies were conducted, and have been labelled Study A (n = 99) and Study B (n = 105). In these studies a subject's judgment profile for a given target (i.e., Richard Hall) was correlated with three different Richard Hall criterion profiles (exemplar, modal profile, and group consensus) to yield three separate sensitivity scores for each subject with respect to the Richard Hall target. This process was repeated for each of the remaining three targets yielding 12 sensitivity scores for each subject. It is the intercorrelations of these 12 scores which are reported in Table 23.

The triangular matrix of correlations for each study can be divided into three groups of correlations, each reporting a different type of generalizability for its sensitivity scores. The first group of correlation coefficients to be considered consists of the four subgroups of three correlations falling on or near the major diagonal of the matrix. Each of these subgroups represents a set of within target, within test, across criterion correlations. As one would expect, in both studies these are generally high with the EX sensitivity-MP sensitivity correlation usually the highest of the three.

The second level of generalizability is represented by the two groups of nine correlations found in the upper left and lower right hand portion of the table for each study. These represent the between target, within test generalizability correlations. For example, the upper left-hand set of nine correlations in Study A and in Study B report the correlations among the three Richard Hall sensitivity scores and the three Ed Murdock sensitivity scores. The lower right-hand portions of the two studies report the correlation between the Jim Anderson sensitivity scores and Jack Cole sensitivity scores. The final block of 36 correlations which is marked off in each table reports the inter-test, inter-target generalizability correlations. Included here are the correlations between the Richard Hall sensitivity scores and the Jim Anderson sensitivity scores, the correlation between the Ed Murdock sensitivity scores and the Jack Cole sensitivity scores and so on, for all possible combinations of PRF and DPI sensitivity scores.

The correlations for each of the studies are nearly all positive and the patterns of magnitudes of correlations within each portion

Table 23

Intercorrelations of Three Sensitivity Measures

for Each of Four Targets

Group A (n = 99)											
	RH EX	RH MP	RH GC	EM EX	EM MP	EM GC	JA EX	JA MP	JA GC	JC EX	JC MP
RH MP	.88										
RH GC	.51	.48									
EM EX	-.01	.14	.11								
EM MP	17	.19	.34	.48							
EM GC	27	32	44	57	.63						
JA EX	23	.13	.26	.03	.18	.40					
JA MP	.23	.19	34	.10	.25	.42	.86				
JA GC	.28	18	.40	02	.22	38	.61	.71			
JC EX	.16	.12	.30	06	.14	.29	.27	.18	.24		
JC MP	07	.06	.27	.12	.16	.28	.23	.17	.17	.92	
JC GC	[illegible]	[illegible]	[illegible]	-.09	.02	.31	.34	.26	.41	.51	.28

Group B (n = 105)											
	RH EX	RH MP	RH GC	EM EX	EM MP	EM GC	JA EX	JA MP	JA GC	JC EX	JC HP
RH MP	.93										
RH GC	.83	.79									
EM EX	10	.20	.19								
EM MP	24	.23	.37	.66							
EM GC	.37	.32	.46	.27	.77						
JA EX	.24	.27	.16	-.11	-.02	.09					
JA MP	27	.30	.22	.08	.23	.26	.82				
JA GC	.38	.34	.37	.08	.31	.42	.33	.48			
JC EX	.34	.32	31	.09	.18	.19	.13	.25	.18		
JC MP	.24	.20	.21	.06	.20	.17	.16	.29	.14	.90	
JC GC	.44	.41	.44	.00	.13	.23	.14	.22	.44	.67	.57

of the table are extremely similar when comparing Study A with Study B. Once can observe that most of the low and the few negative correlations involve the Ed Murdock exemplar scores. This reflects the greater difficulty in identifying good Ed Murdock exemplars from the available sample of college males.

At this point one must change one's perspective in interpreting the strength of these relations. For example one no longer has correlations based on the mean profiles for a large number of subjects. Hence, the reliability of these scores is somewhat lower than is the case for the group consensus profile. Most important is that it seems appropriate to consider each subject's set of sensitivity scores as a set of items measuring general sensitivity. From this perspective it is possible to ascertain the internal consistency of the items and to estimate the reliability of a total sensitivity score derived from the items by computing coefficient alpha. To accomplish this it is necessary to know the average correlation between items (sensitivity scores). To obtain this value the conservative approach was used of including only the 54 inter-target correlations and excluding the 12 intra-target correlations. The resulting mean correlations were .214 for Group A and .231 for Group B. Using Nunnally's Eq. 18-6, coefficient alpha for the 12 sensitivity scores is .766 for Group A, and .783 for Group B. These coefficients indicate the presence of considerable generalizability for the sensitivity parameter.

Implications

In conclusion, one is encouraged by the degree of generality observed in the sensitivity scores. Generality has been demonstrated across both targets and domains of behaviour (tests). Investigators interested in interpersonal perception might take heart in the fact that sensitivity or inferential accuracy can now be assessed with reasonable fidelity, thus facilitating the investigation of relations which may exist between sensitivity and other socially significant variables.

CHAPTER NINE

DISCUSSION

The outcome of these investigations may be summarized as follows:

(1) The average or consensual judgment (group consensus) is very robust in that it generalizes across both judges and items.

(2) The consensual judgment is accurate as assessed with respect to two objective criteria.

(3) The largest component of the consensual judgment profile is clearly associated with the content of items and hence the substance of behaviours, as opposed to other possible components such as base rate of occurrence or desirability.

(4) Individual judge's sensitivities generalize across both targets and domains of behaviour.

These results provide support for the conclusion of the soundness of certain basic ideas underlying this research and other related studies. Some of the ideas were mentioned in Chapter Eight, namely, conceptualizing behaviour in terms of patterns of individual behavioural consistency; using a model of measurement concerned with dimensions of behaviour; and representing the patterns of consistency in terms of profiles of behaviour probabilities. In addition, the strength and consistency of the relations reported here strongly indicate the usefulness of the concept of a multidimensional network of relations among behaviours.

The function of this network in the process of making inferences or in predicting the behaviour of others has been presented as a central feature of the model for inferential accuracy (Jackson, 1972; Reed & Jackson, 1975). Accordingly, the appropriateness of this model is substantiated by these results, particularly with respect to the assessment of accuracy (sensitivity).

Some comments are in order at this point regarding how the individual judgments are made. When a judge is presented with behavioural information about a target it is suggested that this information is used by the judge to get a fix on a location in the inferential network. Then, on the basis of the relative distances among behaviours as well as their relative distance from this initial location, it is possible for the judge to make inferences as to the likelihood that the target under consideration would engage in other specific behaviours. Of course, this is only a model and by no means is it implied that judges have vivid images of multidimensional configurations of behaviours which are referred to as each judgment is made. This is not the purpose of a model.

The usefulness of the model is determined, as is the case for all models, by its ability to achieve two ends. The first, which might be called its predictive value, is the extent to which it accounts for variation in real data. The predictive value of the model for inferential accuracy is evident from the data presented in Chapter Eight, and from other studies such as Reed and Jackson (1975) and Paquin and Jackson (1976).

The second basis on which a model can be evaluted is the degree to which it offers a meaningful conceptualization of the process it is attempting to model. This could be called the model's theoretical value. The theoretical value of the model for inferential accuracy is claimed on the strength of two factors. First is the analogy drawn between Euclidean distances among points in n-dimensional space, and psychological distance among stimuli in an n-dimensional configuration. The soundness of this analogy appears to rest securely on the work of people such as Torgerson (1952, 1958), Messick (1956), Jackson (1962, 1969), and Rosenberg (Rosenberg, Nelson & Vivekananthan, 1968; Rosenberg & Sedlak, 1972). The second factor supporting the theoretical value of the model is the extent to which its parameters represent psychologically meaningful aspects of the judgment process while simultaneously comprising a mathematically sound representation of similarity between judged and criterion profiles. These aspects of the model were developed in Chapters Six and Seven.

The Direction of Future Research

One of the primary purposes of the empirical investigations reported in Chapter Eight was to examine the group consensus judgment with respect to the appropriateness of using it as a criterion for accuracy. This examination was undertaken for two reasons. One had to do with theoretical issues which have been discussed at length above. The second reason was a desire to provide guidelines for the selection of criteria. Prior to this study, the group consensus had been used exclusively as the criterion for assessing the sensitivity of the individual judge (Burron & Jackson, 1974; Jackson, 1972; Paquin & Jackson, 1976; Reed & Jackson, 1975; Strasburger & Jackson, 1975). On the basis of the data reported here it is possible to determine the extent to which this practice is legitimate and advisable.

Concerning the legitimacy of the practice there is clearly no problem. The validity or accuracy of the group consensus judgment is more than evident from the studies reported here. The only question that arises is whether the group consensus is the best single criterion. It seems apparent that in most instances it is not. The best criterion appears to be the modal profile for the group of individuals the target is meant to represent. While in some respects one would think an exemplar profile would be best, the present study indicates that sometimes a good set of exemplars for a specific target may not be available. When this is the case, one could be misled as to the accuracy being achieved, as was true for the Ed Murdock judgments in the present research. Therefore, as a general rule the modal profile seems the most advisable single criterion.

The use of a single exemplar might also seem appealing as a criterion. The problem here is criterion reliability. A single profile is simply not as reliable as the mean of several, which in turn is not as reliable as the factor score for many. As a consequence, while a single exemplar profile may appear more real in one sense, from a measurement point of view its use dooms one to attenuated accuracy correlations due to the lower reliability of this criterion.

Similarly, the use of more than one measure of sensitivity for each subject seems advisable. In fact, the best overall strategy would be to use all three criteria (modal profile, exemplar, and group consensus) and to use the mean of the three resulting sensitivity scores.

There are of course some instances in which neither a factor score (modal profile) nor a group of exemplars is available. For example, consider the situation in which one is interested in comparing the levels of accuracy achieved by subjects judging typical versus atypical targets. In this instance atypicality would be defined in terms of there being a small number (preferably one) of exemplars for the target. Naturally this would also dictate that a meaningful factor score would not be available for the target either. Thus, one would be limited to either a single profile or the group consensus profile as a criterion for assessing individual judgmental accuracy. In this case one would probably opt for the group consensus, assuming it was based on a sufficient number of judges whose judgments have not been contaminated by an experimental manipulation. The group consensus is the choice based on the evidence of its accuracy and reliability reported above.

If one is in the situation of using the group consensus, the interesting possibility exists of improving it as a criterion by removing the desirability variance. In considering this alternative one should keep in mind the case of Richard Hall, where a large overlapping of content and desirability variance made it impossible to separate the two. Thus, while Investigation II of this research suggests that the identification and removal of sources of bias in the group consensus is a possibility, this strategy can produce undesirable results under certain conditions and therefore should be used judiciously.

Obviously these suggestions for using the model for inferential accuracy reveal a degree of expectation that the model will be adopted

Jackson, D. N. The dynamics of structured personality tests: 1971. Psychological Review, 1971, 78, 229-248.

Jackson, D. N. A model for inferential accuracy. The Canadian Psychologist, 1972, 13, 185-194.

Jackson, D. N. Personality Research Form Manual. Rev. Ed. Goshen, N. Y.: Research Psychologists Press, 1974.

Jackson, D. N. Discriminantly valid personality measures: Some propositions. Paper presented at a symposium on the "Future of Personality Assessment" at the meeting of the American Psychological Association, Chicago, August 30, 1975.

Jackson, D. N., & Carlson, K. A. Convergent and discriminant validation of the Differential Personality Inventory. Journal of Clinical Psychology, 1973, 29, 214-219.

Jackson, D. N., & Messick, S. Acquiescence and desirability as response determinants on the MMPI. Educational and Psychological Measurement, 1961, 21, 771-790.

Jackson, D. N., & Messick, D. N. Response styles and the assessment of psychopathology. In D. N. Jackson and S. Messick (Eds.), Problems in human assessment. New York: McGraw-Hill, 1967.

Jackson, D. N., & Messick, S. Differential Personality Inventory. London, Ontario: Authors, 1970.

Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. Attribution: Perceiving the causes of behavior. Morristown, N. J.: General Learning Press, 1972.

References

Allport, G. W. Traits revisited. American Psychologist, 1966, 21, 1-10.

Anastasi, A. Psychological testing. New York: MacMillan, 1976.

Asch, S. E. Forming impressions of personality. Journal of Abnormal and Social Psychology, 1946, 41, 258-290.

Bandura, A. Social learning theory. New York: General Learning Press, 1971.

Bentler, P. M. Response variability: Fact or artifact? Unpublished doctoral dissertation, Stanford University, 1964.

Bowers, K. S. Situationism in psychology: An analysis and a critique. Psychological Review, 1973, 80, 307-336.

Bruner, J. S., & Tagiuri, R. The perception of people. In G. Lindzey (Ed.), Handbook of social psychology. Vol. 2. Cambridge, Mass.: Addison-Wesley, 1954.

Burron, B. F., & Jackson, D. N. Inferential accuracy and threshold in person perception. Research Bulletin No. 296, Department of Psychology, The University of Western Ontario, 1974.

Cattell, R. B. The description and measurement of personality. Yonkers-on-Hudson, New York: World Book, 1946.

Cattell, R. B. Personality and motivation structure and measurement. Yonkers-on-Hudson, N. Y.: World Book, 1957.

Chapman, L. J., & Chapman, J. P. Associatively based illusory correlation as a source of psychodiagnostic folklore. In L. D. Goodstein and R. I. Lanyon (Eds.), Readings in personality assessment. New York: Wiley, 1971, 558-589.

Cline, V. B., & Richards, J. M. Accuracy of interpersonal perception: A general trait? Journal of Abnormal and Social Psychology, 1960, 60, 1-7.

Cronbach, L. J. Processes affecting scores on "understanding of others" and "assumed similarity." Psychological Bulletin, 1955, 52, 177-193.

Cronbach, L. J., & Gleser, G. C. Assessing similarity between profiles. Psychological Bulletin, 1953, 50, 456-473.

Cronbach, L. J., & Meehl, P. E. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.

Crow, W. J., & Hammond, K. R. The generality of accuracy and response sets in interpersonal perception. Journal of Abnormal and Social Psychology, 1957, 54, 384-390.

D'Andrade, R. G. Trait psychology and componential analysis. American Anthropologist, 1965, 67, 215-228.

Fishbein, M., & Ajzen, I. Attitudes toward objects as predictors of single and multiple behavioral criteria. Psychological Review, 1974, 81, 59-74.

Fiske, D. W. Measuring the concepts of personality. Chicago: Aldine, 1971.

Goldberg, L. R. Diagnosticians vs. diagnostic signs: The diagnosis of psychosis vs. neurosis from the MMPI. Psychological Monographs, 1965, 79 (9, Whole no. 602).

Goldberg, L. R. Simple models or simple processes: Some research on clinical judgments. American Psychologist, 1968, 23, 483-496.

Goldberg, L. R. The search for configural relationships in personality assessment. The diagnosis of psychosis vs. neurosis from the MMPI. Multivariate Behavioral Research, 1969, 4, 523-536.

Goldberg, L. R. Man versus model of man: A rationale plus evidence for a method of improving on clinical inferences. Psychological Bulletin, 1970, 73, 422-432.

Golden, M. Some effects of combining psychological tests on clinical inferences. Journal of Consulting Psychology, 1964, 28, 440-446.

Grebstein, L. Relative accuracy of actuarial prediction, experienced clinicians, and graduate students in a clinical judgment task. Journal of Consulting Psychology, 1963, 37, 127-132.

Green, B. F. Descriptions and explanations: A comment on papers by Hoffman and Edwards. In B. Kleinmuntz (Ed.), Formal representation of human judgment. New York: Wiley, 1968.

Harman, H. H. Modern factor analysis. Rev. Ed. Chicago: University of Chicago Press, 1967.

Hastorf, A. H., Schneider, D. J., & Polefka, J. Person perception. Reading, Mass.: Addison-Wesley, 1970.

Horn, J. L. Factor analysis with variables of different metric. Educational and Psychological Measurement, 1969, 29, 753-762.

Hursch, C. J., Hammond, K. R., & Hursch, J. L. Some methodological consideration in multiple-cue probability studies. Psychological Review, 1964, 71, 42-60.

Jackson, D. N. The measurement of perceived personality trait relationships. In Washburne (Ed.), Decisions, values, and groups. Oxford: Pergamon Press, 1962.

Jackson, D. N. Personality Research Form Manual. Goshen, N. Y.: Research Psychologists Press, 1967.

Jackson, D. N. The multidimensional nature of human social perception. Canadian Journal of Behavioural Science, 1969, 1, 227-262.

Jackson, D. N. A sequential system for personality scale development. In C. D. Spielberger (Ed.), Current topics in clinical and community psychology, Vol. 2. New York: Academic Press, 1970.

Jackson, D. N. The dynamics of structured personality tests: 1971. Psychological Review, 1971, 78, 229-248.

Jackson, D. N. A model for inferential accuracy. The Canadian Psychologist, 1972, 13, 185-194.

Jackson, D. N. Personality Research Form Manual. Rev. Ed. Goshen, N. Y.: Research Psychologists Press, 1974.

Jackson, D. N. Discriminantly valid personality measures: Some propositions. Paper presented at a symposium on the "Future of Personality Assessment" at the meeting of the American Psychological Association, Chicago, August 30, 1975.

Jackson, D. N., & Carlson, K. A. Convergent and discriminant validation of the Differential Personality Inventory. Journal of Clinical Psychology, 1973, 29, 214-219.

Jackson, D. N., & Messick, S. Acquiescence and desirability as response determinants on the MMPI. Educational and Psychological Measurement, 1961, 21, 771-790.

Jackson, D. N., & Messick, D. N. Response styles and the assessment of psychopathology. In D. N. Jackson and S. Messick (Eds.), Problems in human assessment. New York: McGraw-Hill, 1967.

Jackson, D. N., & Messick, S. Differential Personality Inventory. London, Ontario: Authors, 1970.

Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., & Weiner, B. Attribution: Perceiving the causes of behavior. Morristown, N. J.: General Learning Press, 1972.

Jones, E. E., & Nisbett, R. E. The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, and B. Weiner (Eds.), Attribution: Perceiving the causes of behavior. Morristown, N. J.: General Learning Press, 1972.

Kahneman, D., & Tversky, A. On the psychology of prediction. Psychological Review, 1973, 80, 237-251.

Kostlan, A. A method for the empirical study of psychodiagnosis. Journal of Consulting Psychology, 1954, 18, 83-88.

Lay, C. H. Trait inferential relationships and judgments about the personality of others. Canadian Journal of Behavioural Science, 1970, 2, 1-17.

Lay, C. H., Burron, B. F., & Jackson, D. N. Base rates and information value in impression formation. Journal of Personality and Social Psychology, 1973, 28, 390-395.

Lay, C. H., & Jackson, D. N. Analysis of the generality of trait inferential relationships. Journal of Personality and Social Psychology, 1969, 12, 12-21.

Lindzey, G. Seer versus sign. Journal of Experimental Research in Personality, 1965, 1, 17-26.

Loevinger, J. Objective tests as instruments of psychological theory. Psychological Reports, 1957, 3, 635-694 (Monograph No. 9).

Lorr, M. Explorations in typing psychotics. Oxford: Pergamon Press, 1966.

Messick, S. J. Some recent developments in multidimensional scaling. Educational and Psychological Measurement, 1956, 16, 82-100.

Mischel, W. Personality and assessment. New York: Wiley, 1968.

Mischel, W. Continuity and change in personality. American Psychologist, 1969, 24, 1012-1018.

Mischel, W. Toward a cognitive social learning reconceptualization of personality. Psychological Review, 1973, 80, 252-283.

Mulaik, S. A. Are personality factors raters' conceptual factors? Journal of Consulting Psychology, 1964, 28, 506-511.

Murray, H. A. Explorations in personality. New York: Oxford University Press, 1938.

Nesselroade, J. R., & Baltes, P. B. Higher order convergence of two distinct personality systems: Cattell's HSPQ and Jackson's PRF. Paper presented at the meetings of the Society of Multivariate Experimental Psychology, Ft. Worth, Texas, November, 1972.

Nisbett, R. E., & Borgida, E. Attribution and the psychology of prediction. Journal of Personality and Social Psychology, 1975, 32, 932-943.

Norman, W. T. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. Journal of Abnormal and Social Psychology, 1963, 66, 574-583.

Norman, W. T., Goldberg, L. R. Raters, ratees, and randomness in personality structure. Journal of Personality and Social Psychology, 1966, 4, 681-691.

Nunnally, J. Psychometric theory. New York, N. Y.: McGraw-Hill, 1967.

Overall, J. E., & Klett, C. J. Applied multivariate analysis. New York: McGraw-Hill, 1972.

Paquin, M. J., & Jackson, D. N. Labelling effects in clinical judgment of psychopathology. Journal of Clinical Psychology, 1976, in press.

Passini, F. T., & Norman, W. T. A universal conception of personality structure. Journal of Personality and Social Psychology, 1966, 4, 44-49.

Reed, P. L., & Jackson, D. N. Clinical judgment of psychopathology: A model for inferential accuracy. Journal of Abnormal Psychology, 1975, 84, 475-482.

Rogers, T. M. The process of responding to personality items: Some issues, a theory and some research. Multivariate Behavioral Research Monographs, 1971, Vol. 6, Number 2.

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. A multidimensional approach to the structure of personality impressions. Journal of Personality and Social Psychology, 1968, 9, 283-294.

Rosenberg, S., & Sedlak, A. Structural representations of implicit personality theory. In L. Berkowitz, (Ed.), Advances in experimental social psychology. Vol. 6. New York: Academic Press, 1972.

Sawyer, J. Measurement and prediction, clinical and statistical. Psychological Bulletin, 1966, 66, 178-200.

Schneider, D. J. Implicit personality theory: A review. Psychological Bulletin, 1973, 79, 294-309.

Sechrest, L., & Jackson, D. N. Social intelligence and accuracy of interpersonal predictions. Journal of Personality, 1961, 29, 167-182.

Sines, L. K. The relative contribution of four kinds of data to accuracy in personality assessment. Journal of Consulting Psychology, 1959, 23, 483-492.

Skinner, H. A. Modal Profile analysis: A unified treatment of classification Procedures. Unpublished doctoral dissertation, The University of Western Ontario, 1975.

Skinner, H. A. Modal profile analysis and classification research. Paper presented at the annual meetings of the Psychometric Society, Murray Hill, New Jersey, April 1-3, 1976.

Skinner, H. A. Classification models for assessing personological variables among Canadian Forces applicants. Paper presented at the annual meetings of the Canadian Psychological Association, Quebec City, Que., June 18-20, 1975. (b).

Skinner, H. A. Differentiating the contribution of elevation, scatter, and shape in profile similarity. Educational and Psychological Measurement, 1977, in press.

Skinner, H. A., & Jackson, D. N. Development of an objective dimensional taxonomy of psychopathology. Paper presented at the meetings of the Society of Multivariate Experimental Psychology, London, Ontario, Canada, November 7-9, 1974.

Skinner, H. A., Jackson, D. N., & Hoffmann, H. Alcoholic personality types: Identification and correlates. Journal of Abnormal Psychology, 1974, 83, 658-666.

Skinner, H. A., Jackson, D. N., & Rampton, G. M. The Personality Research Form in a Canadian context: Does language make a difference? Canadian Journal of Behavioural Science, 1976, 8, 156-168.

Skinner, H. A., Reed, P. L., & Jackson, D. N. Toward the objective diagnosis of psychopathology: Generalizability of Modal Personality profiles. Journal of Consulting and Clinical Psychology, 1976, 44, 111-117.

Strasburger, E. L., & Jackson, D. N. Improving accuracy in a clinical judgment task. Research Bulletin No. 323, Department of Psychology, The University of Western Ontario, 1975.

Stricker, L. J. Personality Research Form: Factor structure and response style involvement. Journal of Consulting and Clinical Psychology, 1974, 42, 529-537.

Stricker, L. J., Jacobs, P. I., & Kogan, N. Trait intercorrelations in implicit personality theories and questionnaire data. Journal of Personality and Social Psychology, 1974, 30, 198-207.

Torgerson, W. S. Multidimensional scaling: I. Theory and method. Psychometrika, 1952, 17, 401-419.

Torgerson, W. S. Theory and methods of scaling. New York: Wiley, 1958.

Tupes, E. C., & Christal, R. E. Recurrent personality factors based on trait ratings. USAF ASD Technical Report, No. 61-97, 1961.

Wiggins, J. S. Personality and prediction: Principles of personality assessment. Reading, Mass.: Addison-Wesley, 1973.

APPENDIX A

Target Descriptions and Instructions

Group Consensus Scale Scores for PRF Targets

Group Consensus Scale Scores for DPI Targets

RICHARD HALL

Richard Hall is a successful young attorney. He has always worked extremely hard and has set high standards and goals for himself. Richard worked tirelessly for long hours in order to attain a high standing in law school. Though finished with his formal education, Richard maintains an interest in several areas of knowledge. He reads many books and his friends think of him as being quite intellectual. Richard entered law because he likes helping people who are having problems. He often gives free legal aid and is always willing to do someone a favor.

Instructions

Read this description two times. Then, using the attached answer sheets, record your judgments of the likelihood that Richard Hall would respond "like" to each of the items in the Statement Booklet. Be sure to print your name and the target name (Richard Hall) at the top of the first answer page.

ED MURDOCK

Ed Murdock lives alone, enjoys doing things his way, and does not like being obligated to other people. He is a loner. At work Ed directs the activities of eleven people. He enjoys being in command and is very forceful with people who disagree with him. Recently he quarrelled with and yelled at someone who made a mistake at work. Ed accused this worker of trying to make him look bad. When questioned by his boss about the way he handled the situation, Ed behaved in a hostile manner and seemed to strongly resent being criticized.

Instructions

Read this description two times. Then, using the attached answer sheets, record your judgments of the likelihood that Ed Murdock would respond "TRUE" to each of the items in the Statement Booklet. Be sure to print your name and the target name (Ed Murdock) at the top of the first answer page.

JIM ANDERSON

Recently Jim Anderson has been feeling very downhearted and "blue." He engages in little activity, and is rarely seen with other people. He has few friends, and seldom talks to the ones he has. Jim holds a very bad opinion of himself and generally thinks he is pretty worthless. When his friends try to talk to him about the way he feels, Jim becomes noticeably uncomfortable and tries to change the subject. In general, they find that Jim is not really aware of any reasons for these feelings and behaviours.

Instructions

Read this description two times. Then, using the attached answer sheets, record your judgments of the likelihood that Jim Anderson would respond "TRUE" to each of the items in the Statement Booklet. Be sure to write your name and the target name (Jim Anderson) at the top of the first answer page.

JACK COLE

Jack Cole has been arrested several times for theft. He says he does not feel guilty about his behaviour and has little feeling for the victims of his thefts. Jack does not appear to worry or care about anything, including being caught. Recently Jack lost a job in a factory because he refused to take his lunch break at the specified time. In his personnel file are recorded several incidents in which he refused to cooperate with his supervisors. In addition, twice it was observed that Jack enjoyed playing rather cruel tricks on fellow workers.

Instructions

Read this description two times. Then, using the attached answer sheets, record your judgments of the likelihood that Jack Cole would respond "TRUE" to each of the items in the Statement Booklet. Be sure to write your name and the target name (Jack Cole) at the top of the first answer page.

Group Concensus (Mean) Scale Scores

for PRF Targets

Scale	Richard Hall APRF	Richard Hall BPRF	Ed Murdock APRF	Ed Murdock BPRF
Abasement	5.50	5.40	2.46	2.68
Achievement	7.28	7.24	5.87	4.56
Affiliation	6.51	6.39	2.99	2.96
Aggression	3.94	4.19	6.81	7.67
Autonomy	3.98	4.50	6.94	7.05
Change	4.93	5.44	4.50	4.13
Cognitive Structure	6.60	6.34	5.29	5.16
Defendence	4.28	4.32	7.23	7.79
Dominance	5.93	6.55	7.04	7.00
Endurance	6.88	7.02	5.03	4.67
Exhibition	4.69	5.47	5.28	4.54
Harmavoidance	5.62	5.53	4.89	4.76
Impulsivity	2.99	2.96	4.89	6.10
Nurturance	6.95	7.17	2.93	3.23
Order	6.55	6.76	5.74	4.85
Play	4.00	4.03	4.39	4.58
Sentience	5.36	5.36	4.46	4.46
Social Recognition	5.50	5.64	4.39	4.94
Succorance	5.35	4.48	3.06	3.11
Understanding	6.38	7.47	4.01	3.64
Infrequency	2.24	2.43	2.84	3.07
Desirability	7.01	6.82	4.96	4.48

Group Concensus (Mean) Scale Scores for

DPI Targets.

Scale	Jim Anderson ADPI	Jim Anderson BDPI	Jack Cole ADPI	Jack Cole BDPI
Broodiness	6.26	7.19	5.84	5.08
Cynicism	5.38	5.39	7.01	7.18
Depression	7.43	7.51	5.36	4.70
Desocialization	7.00	7.46	6.14	5.75
Disorganization of Thinking	5.70	5.48	4.65	4.52
Familial Discord	6.03	5.31	6.96	6.43
Feelings of Unreality	5.78	6.43	4.65	4.66
Health Concern	5.28	5.48	3.55	4.04
Hostility	3.77	3.76	7.49	7.63
Hypochondriasis	5.63	5.58	4.28	4.11
Ideas of Persecution	5.54	5.76	6.26	5.29
Impulsivity	4.42	4.24	6.96	6.69
Irritability	5.89	6.38	6.55	6.34
Mood Fluctuation	5.78	5.26	5.73	4.75
Neurotic Disorganization	5.64	5.45	5.33	5.70
Panic Reaction	6.28	6.61	4.29	3.41
Perceptual Distortion	5.06	4.92	4.03	4.02
Rebelliousness	4.38	4.42	7.55	7.51
Repression	5.17	4.51	3.42	4.38
Sadism	3.88	3.91	6.70	6.32
Self Depreciation	6.34	6.64	4.18	5.10
Shallow Affect	4.05	3.80	5.28	5.98
Socially Deviant Attitudes	3.61	3.76	7.84	7.39
Somatic Complaints	5.72	5.64	4.66	4.16
Defensiveness	4.65	4.76	3.06	3.01
Infrequency	3.47	3.80	3.31	3.16

APPENDIX B

Double Partial Correlations from Investigations I and II

Multiple Correlations and Importance Weights in Predicting the Group Consensus Profile (16-item Scales)

Coefficient Alphas for APRF, BPRF, ADPI, and BDPI

PRF Scale Means for College Students

DPI Scale Means for College Students

PRF Group Consensus Accuracy Correlations for Two Targets:

Population Means and Desirability Partialed Out[a]

	APRF Study n = 99		BPRF Study n = 105	
Criterion Profiles	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile	Richard Hall Group Consensus Profile	Ed Murdock Group Consensus Profile
Hall Exemplars[b]	[illegible] (.70)	.238 (-.07)	.358 (.76)	.146 (-.35)
Hall Modal Profile	[illegible] (.77)	-.017 (-.20)	.419 (.85)	-.070 (-.45)
Murdock Exemplars[b]	-.287 ([illegible])	.332 (.61)	.029 (.28)	.751 (.47)
Murdock Modal Profile	-.486 ([illegible])	.759 (.76)	-.091 (-.19)	.790 (.75)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 21 PRF scales, with the influence of both population means and desirability removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

DPI Group Consensus Accuracy Correlations for Two Targets:

Population Means and Desirability Partialed Out[a]

	ADPI Study n = 99		BDPI Study n = 105	
Criterion Profiles	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile	Jim Anderson Group Consensus Profile	Jack Cole Group Consensus Profile
Anderson Exemplars[b]	.582 (.45)	-.245 (-.50)	.663 (.44)	-.243 (-.30)
Anderson Modal Prof.	.616 (.60)	-.348 (-.54)	.542 (.52)	-.204 (-.30)
Cole Exemplars[b]	-.567 (-.10)	.847 (.79)	-.484 (-.52)	.857 (.78)
Cole Modal Profile	-.553 (-.37)	.673 (.59)	-.474 (-.51)	.787 (.71)

[a]Correlations reflect similarity of group consensus profile and criterion profile shape across 25 DPI scales, with the influence of both population means and desirability removed. Values in parentheses are the original unpartialed correlations.

[b]Exemplar profile for a given target consisted of the mean of 7 male profiles identified as exemplars of the target on the basis of profile shape.

Generalizability of Group Consensus Accuracy for PRF and DPI Targets: Population Means and Desirability Partialed Out

	Generalizability Accuracy r[a] (i.e., APRF Group Consensus Profile: BPRF Exemplar Profile)	Original Accuracy r[b] (i.e., APRF Group Consensus Profile: BPRF Exemplar Profile)
Richard Hall APRF Group Consensus	.018 (.73)[c]	.058 (.70)
Richard Hall BPRF Group Consensus	.156 (.76)	.358 (.83)
Ed Murdock APRF Group Consensus	.751 (.60)	.812 (.61)
Ed Murdock BPRF Group Consensus	.759 (.42)	.751 (.47)
Jim Anderson ADPI Group Consensus	.643 (.42)	.582 (.45)
Jim Anderson BDPI Group Consensus	.511 (.37)	.663 (.44)
Jack Cole ADPI Group Consensus	.698 (.67)	.847 (.79)
Jack Cole BDPI Group Consensus	.851 (.83)	.857 (.78)

[a]Generalizability Accuracy r refers to the correlation of the target group consensus profile from one set of items (i.e., ADPI) with the exemplar profile for the other set of items (BDPI), with the influence of both desirability and and the population means removed.

[b]Original Accuracy r refers to the correlation of the target group consensus and exemplar profiles from the same item set (i.e., ADPI: ADPI), with the influence of both desirability and the population means removed.

[c]Values in parentheses are the unpartialed (zero-order) correlations.

Multiple Correlation and Four Different Types of Importance
Weights in Predicting the Group Consensus Profile:
PRF Targets[a]

Complete Test (16-item scale) - Richard Hall				
$R = .866$			$R^2 = .750$	
	Importance Weights			
Predictor Profiles[b]	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	β
Population Means	.766	.368	.135	-.070
Desirability	.862	.658	.433	.782
Exemplar	.770	.426	.182	.169

Complete Test (16-item scale) - Ed Murdock				
$R = .832$			$R^2 = .693$	
	Importance Weights			
Predictor Profiles	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	β
Population Means	-.310	-.350	.123	-.534
Desirability	-.303	-.329	.108	-.176
Exemplar	.546	.680	.462	.842

[a]It is the group consensus profile that is being predicted in the analyses presented in this table.

[b]Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

Multiple Correlation and Four Different Types of Importance Weights in Predicting the Group Consensus Profile: DPI Targets[a]

Complete Test (16-item scale) - Jim Anderson				
	$R = .741$		$R^2 = .549$	
	Importance Weights			
Predictor Profiles[b]	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	β
Population Means	-.213	-.066	.004	.128
Desirability	-.481	-.535	.286	-.660
Exemplar	.456	.509	.259	.567

Complete Test (16-item scale) - Jack Cole				
	$R = .933$		$R^2 = .871$	
	Importance Weights			
Predictor Profiles[b]	Predictor r	Orthogonal Predictor r	Orthogonal Predictor r^2	β
Population Means	.444	.353	.124	.321
Desirability	-.109	-.341	.116	-.581
Exemplar	.806	.794	.630	.825

[a]It is the group consensus profile that is being predicted in the analyses presented in this table.

[b]Predictor profiles consisted of population scale means, scale means for desirability ratings of items, and an exemplar profile consisting of the mean of 7 males identified as exemplars of the target.

Coefficient Alphas of APRF and BPRF Scales for Inferential Judgment Response Mode (Two Targets)

Scale	Richard Hall A	Richard Hall B	Ed Murdock A	Ed Murdock B
Abasement	.712	.446	.737	.669
Achievement	.511	.537	.809	.671
Affiliation	.794	.702	.834	.653
Aggression	.701	.642	.681	.712
Autonomy	.598	.593	.609	.690
Change	.477	.597	.439	.756
Cognitive Structure	.668	.531	.762	.620
Defendence	.734	.763	.691	.721
Dominance	.642	.628	.729	.611
Endurance	.465	.597	.759	.788
Exhibition	.598	.670	.729	.784
Harmavoidance	.647	.710	.723	.784
Implusivity	.696	.744	.761	.836
Nurturance	.734	.636	.766	.727
Order	.791	.842	.884	.888
Play	.637	.540	.729	.637
Sentience	.604	.571	.562	.475
Social Recognition	.568	.693	.746	.680
Succorance	.631	.659	.693	.681
Understanding	.593	.726	.768	.836
Infrequency	.702	.692	.722	.683
Desirability	.561	.598	.655	.613

Coefficient Alphas of ADPI and BDPI Scales for Inferential Judgment Response Mode (Two Targets)

Scale	Jim Anderson ADPI	Jim Anderson BDPI	Jack Cole ADPI	Jack Cole BDPI
Broodiness	.493	.627	.425	.672
Cynicism	.756	.711	.581	.633
Depression	.868	.705	.842	.665
Desocialization	.833	.655	.779	.830
Disorganization of Thinking	.649	.656	.705	.704
Familial Discord	.769	.773	.785	.675
Feelings of Unreality	.678	.679	.597	.762
Health Concern	.779	.789	.702	.617
Hostility	.661	.742	.675	.785
Hypochondriasis	.662	.789	.685	.777
Ideas of Persecution	.770	.629	.680	.486
Impulsivity	.688	.709	.577	.693
Irritability	.781	.758	.704	.709
Mood Fluctuation	.619	.635	.592	.423
Neurotic Disorganization	.754	.816	.761	.725
Panic Reaction	.686	.796	.738	.708
Perceptual Distortion	.749	.664	.744	.622
Rebelliousness	.760	.511	.628	.500
Repression	.260	.219	.397	.364
Sadism	.760	.698	.662	.580
Self Depreciation	.848	.710	.717	.774
Shallow Affect	.403	.474	.292	.433
Socially Deviant Attitudes	.760	.685	.764	.501
Somatic Complaints	.758	.845	.678	.793
Defensiveness	.668	.290	.569	.570
Infrequency	.710	.662	.718	.714

PRF Scale Means for College Males and Females

(n = 215)

Scale	APRF	BPRF	PRF-E[a]
Abasement	3.79	3.94	7.73
Achievement	5.20	5.28	10.48
Affiliation	4.18	4.21	8.39
Aggression	3.58	3.51	7.09
Autonomy	4.20	4.10	8.30
Change	5.10	4.61	9.71
Cognitive Structure	4.24	4.59	8.84
Defendence	2.81	2.95	5.76
Dominance	4.42	4.36	8.78
Endurance	5.26	5.35	10.61
Exhibition	3.57	3.63	7.20
Harmavoidance	4.31	4.28	8.59
Impulsivity	2.94	2.96	5.90
Nurturance	4.98	5.05	10.03
Order	4.21	4.10	8.31
Play	4.25	4.20	8.45
Sentience	4.95	5.08	10.03
Social Recognition	3.96	3.82	7.78
Succorance	3.67	3.50	7.17
Understanding	5.03	5.03	10.06
Infrequency	0.23	0.21	0.43
Desirability	5.51	5.46	10.97

[a]Due to rounding error, PRF-E column does not always equal sum of APRF and BPRF.

DPI Scale Means for College Males and Females

(n = 179)

Scale	ADPI	BDPI	DPI[a]
Broodiness	2.20	2.85	5.06
Cynicism	2.62	2.01	4.63
Depression	1.38	1.40	2.79
Desocialization	2.22	1.90	4.12
Disorganization of Thinking	0.91	0.96	1.87
Familial Discord	2.35	2.55	4.90
Feelings of Unreality	1.81	1.72	3.53
Health Concern	2.71	2.59	5.30
Hostility	2.38	2.21	4.59
Hypochondriasis	1.19	0.98	2.16
Ideas of Persecution	0.82	1.01	1.83
Impulsivity	3.48	3.10	6.58
Irritability	4.36	4.21	8.56
Mood Fluctuation	3.88	4.13	8.02
Neurotic Disorganization	2.71	2.94	5.64
Panic Reaction	2.31	1.94	4.26
Perceptual Distortion	1.26	1.24	2.50
Rebelliousness	4.80	4.99	9.78
Repression	2.06	1.87	3.93
Sadism	1.70	1.93	3.64
Self Depreciation	0.53	0.61	1.14
Shallow Affect	1.30	1.14	2.44
Socially Deviant Attitudes	2.91	2.98	5.89
Somatic Complaints	1.58	1.50	3.09
Defensiveness	2.55	2.70	5.25
Infrequency	0.15	0.16	0.31

[a]Due to rounding error, DPI column does not always equal sum of ADPI and BDPI.

APPENDIX C

Factor Loadings of Orthogonalized Predictors of Group Consensus Profiles

Intercorrelations of Predictor and Criterion Profiles

Factor Loadings of Orthogonalized Predictors of Group

Consensus Profile: PRF Target Richard Hall

Study A

Predictor Profiles	Factor Loadings I	II	III
Population Means	.804	.417	.425
Desirability	.417	.824	.384
Exemplar	.425	.384	.820

Study B

Predictor Profiles	Factor Loadings I	II	III
Population Means	.792	.446	.416
Desirability	.446	.793	.415
Exemplar	.416	.415	.809

Total Test

Predictor Profiles	Factor Loadings I	II	III
Population Means	.785	.450	.425
Desirability	.450	.794	.409
Exemplar	.425	.409	.807

Factor Loadings of Orthogonalized Predictors of Group

Consensus Profile: PRF Target Ed Murdock

Study A

Predictor Profiles	Factor Loadings I	II	III
Population Means	.845	.474	.246
Desirability	.474	.874	.103
Exemplar	.246	.103	.964

Study B

Predictor Profiles	Factor Loadings I	II	III
Population Means	.840	.508	.192
Desirability	.508	.856	.099
Exemplar	.192	.099	.976

Total Test

Predictor Profiles	Factor Loadings I	II	III
Population Means	.832	.511	.219
Desirability	.511	.853	.112
Exemplar	.219	.112	.969

Factor Loadings of Orthogonalized Predictors of Group Consensus Profiles, DPI Target Jim Anderson

Study A

Predictor Profiles	Factor Loadings I	II	III
Population Means	.951	.299	.082
Desirability	.299	.953	.047
Exemplar	.082	.047	.996

Study B

Predictor Profiles	Factor Loadings I	II	III
Population Means	.950	.311	-.023
Desirability	.311	.943	.118
Exemplar	-.023	.118	.993

Total Test

Predictor Profiles	Factor Loadings I	II	III
Population Means	.951	.307	.027
Desirability	.307	.947	.090
Exemplar	.027	.090	.996

Factor Loadings of Orthogonalized Predictors of Group Consensus Profile: DPI Target Jack Cole

Study A

Predictor Profiles	Factor Loadings I	II	III
Population Means	.905	.291	.310
Desirability	.291	.951	.108
Exemplar	.310	.108	.945

Study B

Predictor Profiles	Factor Loadings I	II	III
Population Means	.930	.293	.223
Desirability	.293	.941	.169
Exemplar	.223	.169	.960

Total Test

Predictor Profiles	Factor Loadings I	II	III
Population Means	.914	.294	.279
Desirability	.294	.946	.138
Exemplar	.279	.138	.950

Intercorrelations of Predictor (Population Mean, Desirability, and Exemplar) and Criterion (Group Consensus) Profiles: PRF Target Richard Hall

Study A

	Population Mean	Desirability	Exemplar
Desirability	.841		
Exemplar	.850	.808	
Group Consensus	.708	.840	.695

Study B

	Population Mean	Desirability	Exemplar
Desirability	.880		
Exemplar	.852	.851	
Group Consensus	.786	.846	.829

Total Test

	Population Mean	Desirability	Exemplar
Desirability	.884		
Exemplar	.861	.847	
Group Consensus	.766	.862	.770

Intercorrelations of Predictor (Population Mean, Desirability, and Exemplar) and Criterion (Group Consensus) Profiles: PRF Target Ed Murdock

Study A

	Population Mean	Desirability	Exemplar
Desirability	.841		
Exemplar	.494	.307	
Group Consensus	-.186	-.249	.607

Study B

	Population Mean	Desirability	Exemplar
Desirability	.880		
Exemplar	.398	.279	
Group Consensus	-.410	-.492	.469

Total Test

	Population Mean	Desirability	Exemplar
Desirability	.884		
Exemplar	.451	.316	
Group Consensus	-.310	-.383	.546

Intercorrelations of Predictor (Population Mean, Desirability, and Exemplar) and Criterion (Group Consensus) Profiles: DPI Target Jim Anderson

Study A

	Population Mean	Desirability	Exemplar
Desirability	.574		
Exemplar	.174	.116	
Group Consensus	-.230	-.490	.450

Study B

	Population Mean	Desirability	Exemplar
Desirability	.586		
Exemplar	-.008	.220	
Group Consensus	-.188	-.461	.443

Total Test

	Population Mean	Desirability	Exemplar
Desirability	.586		
Exemplar	.080	.183	
Group Consensus	-.213	-.481	.456

Intercorrelations of Predictor (Population Mean, Desirability, and Exemplar) and Criterion (Group Consensus) Profiles:

DPI Target Jack Cole

Study A

	Population Mean	Desirability	Exemplar
Desirability	.574		
Exemplar	.605	.295	
Group Consensus	.453	-.174	.790

Study B

	Population Mean	Desirability	Exemplar
Desirability	.586		
Exemplar	.471	.386	
Group Consensus	.397	-.048	.775

Total Test

	Population Mean	Desirability	Exemplar
Desirability	.586		
Exemplar	.561	.343	
Group Consensus	.444	-.109	.806

APPENDIX D

Frequency Distributions and Statistics for Individual Subject Parameters

DISTRIBUTION OF SUBJECT PARAMETERS A-PRE

ELEVATION-HALL

```
I
**************************************** (     8)   8.1 PCT
I    5.15-5.25
I
**************************************** (     8)   8.1 PCT
I    5.25-5.35
I
*************************************************************************** (    15)  15.2 PCT
I    5.35-5.45
I
*********************************************************************************************** (    19)  19.2 PCT
I    5.45-5.55
I
*********************************************************************************************** (    19)  19.2 PCT
I    5.55-5.65
I
******************************************************************************** (    16)  16.2 PCT
I    5.65-5.75
I
************************* (     5)   5.1 PCT
I    5.75-5.85
I
*************** (     3)   3.0 PCT
I    5.85-5.95
I
*************** (     3)   3.0 PCT
I    5.95-6.05
I
*************** (     3)   3.0 PCT
I    6.05-6.15
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         2         4         6         8        10        12        14        16        18        20
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

        ELEVATION-MURDOCK

I
*************** (      7)   7.1 PCT
I     4.25-4.40
I
I
*************** (      7)   7.1 PCT
I     4.40-4.55
I
I
******************* (      9)   9.1 PCT
I     4.55-4.70
I
I
**************************************** (     19)  19.2 PCT
I     4.70-4.85
I
I
************************************************ (     23)  23.2 PCT
I     4.85-5.00
I
I
************************************ (     17)  17.2 PCT
I     5.00-5.15
I
I
**************************** (     13)  13.1 PCT
I     5.15-5.30
I
I
***** (      2)   2.0 PCT
I     5.30-5.45
I
I
*** (      1)   1.0 PCT
I     5.45-5.60
I
I
*** (      1)   1.0 PCT
I     5.60-5.75
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRE

ELEVATION-ANDERSON

```
I
***** (     2)   2.0 PCT
I     3.90-4.15
I
I
*** (     1)   1.0 PCT
I     4.15-4.40
I
I
******* (     3)   3.0 PCT
I     4.40-4.65
I
I
**************** (     6)   6.1 PCT
I     4.65-4.90
I
I
*************************************** (    15)  15.2 PCT
I     4.90-5.15
I
I
***************************************************************** (    25)  25.3 PCT
I     5.15-5.40
I
I
********************************************* (    17)  17.2 PCT
I     5.40-5.65
I
I
****************************************** (    16)  16.2 PCT
I     5.65-5.90
I
I
****************** (     7)   7.1 PCT
I     5.90-6.15
I
I
**************** (     6)   6.1 PCT
I     6.15-6.40
I
I
*** (     1)   1.0 PCT
I     6.40-6.65
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRP

ELEVATION-COLE

```
I
****  (     2)   2.0 PCT
I    4.50-4.70
I
I
**********  (     5)   5.1 PCT
I    4.70-4.90
I
I
****************  (     8)   8.1 PCT
I    4.90-5.10
I
I
****************************  (    14)  14.1 PCT
I    5.10-5.30
I
I
******************************************  (    21)  21.2 PCT
I    5.30-5.50
I
I
**********************************  (    17)  17.2 PCT
I    5.50-5.70
I
I
****************************  (    14)  14.1 PCT
I    5.70-5.90
I
I
********************  (    10)  10.1 PCT
I    5.90-6.10
I
I
******  (     3)   3.0 PCT
I    6.10-6.30
I
I
********  (     4)   4.0 PCT
I    6.30-6.50
I
I
**  (     1)   1.0 PCT
I    6.50-6.70
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5         10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS H-PRR

ELEVATION=[illegible]LL

Interval	Frequency	Percent
4.75-4.95	1	1.0 PCT
4.95-5.15	1	1.0 PCT
5.15-5.35	8	7.6 PCT
5.35-5.55	25	23.8 PCT
5.55-5.75	30	28.6 PCT
5.75-5.95	29	27.6 PCT
5.95-6.15	6	5.7 PCT
6.15-6.35	3	2.9 PCT
6.35-6.55	2	1.9 PCT

FREQUENCY: 0 5 10 15 20 25 30 35 40 45 50

DISTRIBUTION OF SUBJECT PARAMETERS H-RRF

ELEVATION=MURDOCK

```
I
******** (     4)   3.8 PCT
I    3.90-4.10
I
I
**** (     2)   1.9 PCT
I    4.10-4.30
I
I
************** (     7)   6.7 PCT
I    4.30-4.50
I
I
************************** (    13)  12.4 PCT
I    4.50-4.70
I
I
****************************************************** (    27)  25.7 PCT
I    4.70-4.90
I
I
****************************************************** (    27)  25.7 PCT
I    4.90-5.10
I
I
****************************** (    15)  14.3 PCT
I    5.10-5.30
I
I
************** (     7)   6.7 PCT
I    5.30-5.50
I
I
****** (     3)   2.9 PCT
I    5.50-5.70
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS H-POF

ELEVATION-ANDERSON

```
I
************** (    7)   6.7 PCT
I    4.55-4.80
I
I
********************************** (    17)  16.2 PCT
I    4.80-5.05
I
I
********************************************** (    23)  21.9 PCT
I    5.05-5.30
I
I
******************************** (    16)  15.2 PCT
I    5.30-5.55
I
I
**************************************** (    20)  19.0 PCT
I    5.55-5.80
I
I
****************************** (    15)  14.3 PCT
I    5.80-6.05
I
I
******** (    4)   3.8 PCT
I    6.05-6.30
I
I
**** (    2)   1.9 PCT
I    6.30-6.55
I
I
** (    1)   1.0 PCT
I    6.55-6.80
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS H-PPE

          ELEVATION-COLE

I
*** (     1)   1.0 PCT
I   4.00-4.25
I
*** (     1)   1.0 PCT
I   4.25-4.50
I
************* (     6)   5.7 PCT
I   4.50-4.75
I
********************** (    10)   9.5 PCT
I   4.75-5.00
I
********************************************** (    22)  21.0 PCT
I   5.00-5.25
I
***************************************************** (    25)  23.8 PCT
I   5.25-5.50
I
*************************************************** (    24)  22.9 PCT
I   5.50-5.75
I
******************** (     9)   8.6 PCT
I   5.75-6.00
I
*********** (     5)   4.8 PCT
I   6.00-6.25
I
***** (     2)   1.9 PCT
I   6.25-6.50
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5         10        15        20        25        30        35        40        45        50
FREQUENCY
```

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - AFRF

ELEVATION

HALL

MEAN	4.566	STD ERROR	.214	MEDIAN	4.474
MODE	5.000	STD DEV	2.134	VARIANCE	4.554
KURTOSIS	.066	SKEWNESS	.450	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

MURDOCK

MEAN	4.717	STD ERROR	.190	MEDIAN	4.826
MODE	5.000	STD DEV	1.890	VARIANCE	3.572
KURTOSIS	-.134	SKEWNESS	-.098	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

ANDERSON

MEAN	6.495	STD ERROR	.196	MEDIAN	6.400
MODE	6.000	STD DEV	1.950	VARIANCE	3.804
KURTOSIS	.324	SKEWNESS	-.250	RANGE	10.000
MINIMUM	1.000	MAXIMUM	11.000		

COLE

MEAN	5.606	STD ERROR	.211	MEDIAN	5.476
MODE	5.000	STD DEV	2.104	VARIANCE	4.425
KURTOSIS	-.695	SKEWNESS	.171	RANGE	10.000
MINIMUM	1.000	MAXIMUM	11.000		

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - BRRF
ELEVATION

HALL

MEAN	5.095	STD ERROR	.133	MEDIAN	5.083
MODE	5.000	STD DEV	1.362	VARIANCE	1.856
KURTOSIS	.776	SKEWNESS	.194	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

MURDOCK

MEAN	5.390	STD ERROR	.167	MEDIAN	5.481
MODE	5.000	STD DEV	1.712	VARIANCE	2.933
KURTOSIS	.349	SKEWNESS	-.813	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

ANDERSON

MEAN	3.971	STD ERROR	.173	MEDIAN	3.844
MODE	3.000	STD DEV	1.773	VARIANCE	3.143
KURTOSIS	-.444	SKEWNESS	.314	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

COLE

MEAN	5.962	STD ERROR	.164	MEDIAN	6.000
MODE	6.000	STD DEV	1.681	VARIANCE	2.825
KURTOSIS	.149	SKEWNESS	-.123	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

A

DISPERSION-HALL

Interval	Frequency	Percent
0.70-0.85	7	7.1 PCT
0.85-1.00	2	2.0 PCT
1.00-1.15	13	13.1 PCT
1.15-1.30	11	11.1 PCT
1.30-1.45	14	14.1 PCT
1.45-1.60	17	17.2 PCT
1.60-1.75	13	13.1 PCT
1.75-1.90	19	19.2 PCT
1.90-2.05	5	5.0 PCT

0 2 4 6 8 10 12 14 16 18 20

FREQUENCY

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF.

DISPERSION-MURDOCK

Range	Frequency	Percent
0.65-0.85	1	1.0 PCT
0.85-1.05	5	5.1 PCT
1.05-1.25	9	9.1 PCT
1.25-1.45	12	12.1 PCT
1.45-1.65	14	14.1 PCT
1.65-1.85	20	20.2 PCT
1.85-2.05	19	19.2 PCT
2.05-2.25	12	12.1 PCT
2.25-2.45	4	4.0 PCT
2.45-2.65	3	3.0 PCT

FREQUENCY: 0 2 4 6 8 10 12 14 16 18 20

```
DISTRIBUTION OF SUBJECT PARAMETERS, A-PRE

        DISPERSION-ANDERSON

I
*** (      1)   1.0 PCT
I    0.00-0.10
I
I
************* (      6)   6.1 PCT
I    0.10-0.20
I
I
***************** (      8)   8.1 PCT
I    0.20-0.30
I
I
*********************************** (     17)  17.2 PCT
I    0.30-0.40
I
I
***************************************** (     20)  20.2 PCT
I    0.40-0.50
I
I
********************************************************* (     28)  28.3 PCT
I    0.50-0.60
I
I
********************* (     10)  10.1 PCT
I    0.60-0.70
I
I
*********** (      5)   5.1 PCT
I    0.70-0.80
I
I
******* (      3)   3.0 PCT
I    0.80-0.90
I
I
*** (      1)   1.0 PCT
I    0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

DISPERSION-COLE

Range	Frequency	Percent
0.40-0.65	2	2.0 PCT
0.65-0.90	4	4.0 PCT
0.90-1.15	8	8.0 PCT
1.15-1.40	17	17.2 PCT
1.40-1.65	22	22.2 PCT
1.65-1.90	30	30.3 PCT
1.90-2.15	17	17.2 PCT
2.15-2.40	2	2.0 PCT
2.65-2.90	1	1.0 PCT

FREQUENCY 0 5 10 15 20 25 30 35 40 45 50

DISTRIBUTION OF SUBJECT PARAMETERS H-PRE

DISPERSION-HALL

Interval	Frequency	Percent
0.20-0.40	1	4.0 PCT
0.40-0.60	1	1.0 PCT
0.60-0.80	4	3.8 PCT
0.80-1.00	8	7.6 PCT
1.00-1.20	15	14.3 PCT
1.20-1.40	18	17.1 PCT
1.40-1.60	19	18.1 PCT
1.60-1.80	16	15.2 PCT
1.80-2.00	14	13.3 PCT
2.00-2.20	9	8.6 PCT

FREQUENCY 0 2 4 6 8 10 12 14 16 18 20

DISTRIBUTION OF SUBJECT PARAMETERS B-PRE

DISPERSION-FURROCK

```
****     (   2)   1.9 PCT
I  0.25-0.50
I
**       (   1)   1.0 PCT
I
I
************ (   6)   5.7 PCT
I  0.75-1.00
I
**********   (   5)   4.8 PCT
I  1.00-1.25
I
**************************   (  13)  12.4 PCT
I  1.25-1.50
I
********************************************   (  22)  21.0 PCT
I  1.50-1.75
I
**************************************   (  19)  18.1 PCT
I  1.75-2.00
I
******************************************************   (  27)  25.7 PCT
I  2.00-2.25
I
**************   (   7)   6.7 PCT
I  2.25-2.50
I
******   (   3)   2.9 PCT
I  2.50-2.75
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS H-PRF,

          DISPERSION-ANDERSON

I
*** (       1)    1.0 PCT
I     0.15-0.40
I
I
******* (       3)    2.9 PCT
I     0.40-0.65
I
I
*********** (       5)    4.8 PCT
I     0.65-0.90
I
I
***************************** (      14)   13.3 PCT
I     0.90-1.15
I
I
***************************************************** (      26)   24.8 PCT
I     1.15-1.40
I
I
***************************************************************** (      32)   30.[illegible] PCT
I     1.40-1.65
I
I
********************************* (      16)   15.2 PCT
I     1.65-1.90
I
I
*************** (       7)    6.7 PCT
I     1.90-2.15
I
I
*** (       1)    1.0 PCT
I     2.15-2.40
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS N=PR

DISPERSION=COLE

I
*** (      1)   1.0 PCT
I    0.25-0.50
I
I
********* (      4)   3.8 PCT
I    0.50-0.75
I
*********** (      5)   4.8 PCT
I    0.75-1.00
I
I
************************************ (     17)  16.2 PCT
I    1.00-1.25
I
I
******************************************* (     20)  19.0 PCT
I    1.25-1.50
I
I
********************************************* (     21)  20.0 PCT
I    1.50-1.75
I
I
****************************************** (     19)  18.1 PCT
I    1.75-2.00
I
I
********************************* (     15)  14.3 PCT
I    2.00-2.25
I
I
***** (      2)   1.9 PCT
I    2.25-2.50
I
I
*** (      1)   1.0 PCT
I    2.50-2.75
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - APRF

DISPERSION

HALL

MEAN	5.014	STD ERROR	.219	MEDIAN	5.607
MODE	8.000	STD DEV	2.176	VARIANCE	4.735
KURTOSIS	-.772	SKEWNESS	-.373	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

MURDOCK

MEAN	5.768	STD ERROR	.201	MEDIAN	5.925
MODE	6.000	STD DEV	2.004	VARIANCE	4.017
KURTOSIS	-.499	SKEWNESS	-.137	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

ANDERSON

MEAN	5.253	STD ERROR	.175	MEDIAN	5.375
MODE	6.000	STD DEV	1.746	VARIANCE	3.048
KURTOSIS	.033	SKEWNESS	.048	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

COLE

MEAN	5.303	STD ERROR	.155	MEDIAN	5.517
MODE	6.000	STD DEV	1.542	VARIANCE	2.377
KURTOSIS	.767	SKEWNESS	-.432	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - BPRF

DISPERSION

HALL

MEAN	6.733	STD ERROR	.196	MEDIAN	6.789
MODE	7.000	STD DEV	2.011	VARIANCE	4.044
KURTOSIS	-.398	SKEWNESS	-.287	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

MURDOCK

MEAN	6.486	STD ERROR	.183	MEDIAN	6.684
MODE	8.000	STD DEV	1.877	VARIANCE	3.521
KURTOSIS	.318	SKEWNESS	-.671	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

ANDERSON

MEAN	5.495	STD ERROR	.144	MEDIAN	5.600
MODE	6.000	STD DEV	1.475	VARIANCE	2.175
KURTOSIS	.398	SKEWNESS	-.406	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

COLE

MEAN	5.705	STD ERROR	.173	MEDIAN	5.762
MODE	6.000	STD DEV	1.770	VARIANCE	3.133
KURTOSIS	-.324	SKEWNESS	-.215	RANGE	9.000
MINIMUM	1.000	MAXIMUM	10.000		

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

          MODAL PROFILE SENSITIVITY-HALL

I
********* (       4)   4.0 PCT
I    0.30-0.40
I
I
************* (       6)   6.1 PCT
I    0.40-0.50
I
I
****************************************************** (     27)  27.3 PCT
I    0.50-0.60
I
I
************************************************************** (     31)  31.3 PCT
I    0.60-0.70
I
I
****************************************************** (     27)  27.3 PCT
I    0.70-0.80
I
I
********* (       4)   4.0 PCT
I    0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

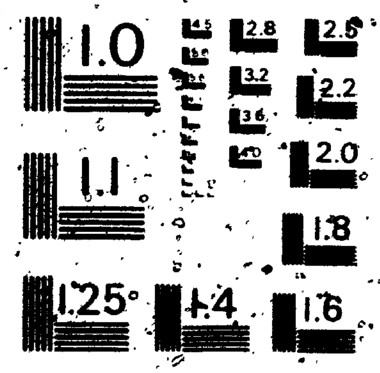
```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

          MODAL PROFILE SENSITIVITY-MURDOCK

I
*** (     1)   1.0 PCT
I    -1.00-0.00
I
I
***** (     2)   2.0 PCT
I    0.20-0.30
I
I
********* (     4)   4.0 PCT
I    0.30-0.40
I
I
********************** (    10)  10.1 PCT
I    0.40-0.50
I
I
************************************ (    17)  17.2 PCT
I    0.50-0.60
I
I
************************************************************ (    29)  29.3 PCT
I    0.60-0.70
I
I
************************************************************ (    29)  29.3 PCT
I    0.70-0.80
I
I
*************** (     7)   7.1 PCT
I    0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

MODAL PROFILE SENSITIVITY ANDERSON

```
I
************ (      5)   5.1 PCT
I    -1.00-0.00
I
************ (      5)   5.1 PCT
I    0.00-0.10
I
*** (      1)   1.0 PCT
I    0.10-0.20
I
************** (      7)   7.1 PCT
I    0.20-0.30
I
********************************************* (     22)  22.2 PCT
I    0.30-0.40
I
********************************************* (     22)  22.2 PCT
I    0.40-0.50
I
*********************************** (     16)  16.2 PCT
I    0.50-0.60
I
************************* (     12)  12.1 PCT
I    0.60-0.70
I
******************* (      9)   9.1 PCT
I    0.70-0.80
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
          5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

3 OF/DE 3



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

DISTRIBUTION OF SUBJECT PARAMETERS A-PHF

MODAL PROFILE SENSITIVITY-COLE

```
I
***** (      2)   2.0 PCT
I     0.00-0.10
I
***** (      2)   2.0 PCT
I     0.10-0.20
I
************************ (     12)  12.1 PCT
I     0.20-0.30
I
**************** (      8)   8.1 PCT
I     0.30-0.40
I
****************************************** (     19)  19.2 PCT
I     0.40-0.50
I
************************************************* (     22)  22.2 PCT
I     0.50-0.60
I
************************************************************** (     31)  31.3 PCT
I     0.60-0.70
I
***** (      2)   2.0 PCT
I     0.70-0.80
I
*** (      1)   1.0 PCT
I     0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS H-PHF

            MODAL PROFILE SENSITIVITY-HALL

I
*** (      1)    1.0 PCT
I     -1.00-0.00
I
I
******* (      3)    2.9 PCT
I      0.30-0.40
I
I
*************** (      7)    6.7 PCT
I      0.40-0.50
I
I
*************** (      7)    6.7 PCT
I      0.50-0.60
I
I
******************************************** (     22)   21.0 PCT
I      0.60-0.70
I
I
**************************************************************************************** (     44)   41.9 PCT
I      0.70-0.80
I
I
************************************ (     18)   17.1 PCT
I      0.80-0.90
I
I
******* (      3)    2.9 PCT
I      0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS R-PBF

MODAL PROFILE SENSITIVITY-MURDOCK

-1.00-0.00	(1)	1.0 PCT
0.10-0.20	(2)	1.9 PCT
0.20-0.30	(3)	2.9 PCT
0.30-0.40	(4)	3.8 PCT
0.40-0.50	(9)	8.6 PCT
0.50-0.60	(21)	20.0 PCT
0.60-0.70	(30)	28.6 PCT
0.70-0.80	(27)	25.7 PCT
0.80-0.90	(8)	7.6 PCT

0 5 10 15 20 25 30 35 40 45 50

FREQUENCY

DISTRIBUTION OF SUBJECT PARAMETERS H-PRE

MODAL PROFILE SENSITIVITY-ANDERSON

Range	Frequency	Percent
-1.00-0.00	2	1.9 PCT
0.00-0.10	1	1.0 PCT
0.10-0.20	12	11.4 PCT
0.20-0.30	13	12.4 PCT
0.30-0.40	[illegible]	17.1 PCT
0.40-0.50	24	22.9 PCT
0.50-0.60	21	20.0 PCT
0.60-0.70	9	8.6 PCT
0.70-0.80	5	4.8 PCT

FREQUENCY (axis: 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50)

DISTRIBUTION OF SUBJECT PARAMETERS H-PRF

MODAL PROFILE SENSITIVITY-COLE

Interval	Frequency	Percent
-1.00-0.00	1	1.0 PCT
0.00-0.10	1	1.0 PCT
0.20-0.30	6	5.7 PCT
0.30-0.40	3	2.8 PCT
0.40-0.50	9	8.6 PCT
0.50-0.60	33	31.4 PCT
0.60-0.70	33	31.4 PCT
0.70-0.80	18	17.1 PCT
0.80-0.90	1	1.0 PCT

0 5 10 15 20 25 30 35 40 45 50

FREQUENCY

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - APRP

MODAL PROFILE SENSITIVITY

HALL

MEAN	6.830	STD ERROR	.115	MEDIAN	6.903
MODE	7.000	STD DEV	1.140	VARIANCE	1.300
KURTOSIS	[illegible].005	SKEWNESS	-.420	RANGE	5.000
MINIMUM	4.000	MAXIMUM	9.000		

MURDOCK

MEAN	6.788	STD ERROR	.153	MEDIAN	7.034
MODE	8.000	STD DEV	1.520	VARIANCE	2.312
KURTOSIS	3.009	SKEWNESS	-1.301	RANGE	9.000
MINIMUM	0.000	MAXIMUM	9.000		

ANDERSON

MEAN	4.828	STD ERROR	.205	MEDIAN	4.932
MODE	4.000	STD DEV	2.036	VARIANCE	4.144
KURTOSIS	.093	SKEWNESS	-.590	RANGE	8.000
MINIMUM	0.000	MAXIMUM	8.000		

COLE

MEAN	5.685	STD ERROR	.165	MEDIAN	5.795
MODE	7.000	STD DEV	1.637	VARIANCE	2.681
KURTOSIS	-.133	SKEWNESS	-.681	RANGE	8.000
MINIMUM	1.000	MAXIMUM	9.000		

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - BPRF

MODAL-PROFILE SENSITIVITY

HALL

MEAN	7.495	STD ERROR	.105	MEDIAN	7.784
MODE	8.000	STD DEV	1.481	VARIANCE	2.195
KURTOSIS	5.313	SKEWNESS	-1.693	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

MURDOCK

MEAN	6.648	STD ERROR	.162	MEDIAN	6.917
MODE	7.000	STD DEV	1.658	VARIANCE	2.750
KURTOSIS	2.161	SKEWNESS	-1.246	RANGE	9.000
MINIMUM	0.000	MAXIMUM	9.000		

ANDERSON

MEAN	4.619	STD ERROR	.174	MEDIAN	4.771
MODE	5.000	STD DEV	1.783	VARIANCE	3.180
KURTOSIS	-.353	SKEWNESS	-.252	RANGE	8.000
MINIMUM	0.000	MAXIMUM	8.000		

COLE

MEAN	6.267	STD ERROR	.149	MEDIAN	6.485
MODE	6.000	STD DEV	1.527	VARIANCE	2.332
KURTOSIS	2.936	SKEWNESS	-1.432	RANGE	9.000
MINIMUM	0.000	MAXIMUM	9.000		

DISTRIBUTION OF SUBJECT PARAMETERS A-PRE

EXEMPLAR SENSITIVITY-HALL

```
I
I** (      1)   1.0 PCT
I    0.20-0.30
I
I************ (      6)   6.1 PCT
I    0.30-0.40
I
I****************************** (     15)  15.2 PCT
I    0.40-0.50
I
I************************************************************** (     31)  31.3 PCT
I    0.50-0.60
I
I************************************************************ (     30)  30.3 PCT
I    0.60-0.70
I
I**************************** (     14)  14.1 PCT
I    0.70-0.80
I
I**** (      2)   2.0 PCT
I    0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

        EXEMPLAR SENSITIVITY-MURDOCK

I
*** (     1)   1.0 PCT
I   -1.00-0.00
I
I
********* (     4)   4.0 PCT
I    0.00-0.10
I
I
******* (     3)   3.0 PCT
I    0.10-0.20
I
I
***************** (     8)   8.1 PCT
I    0.20-0.30
I
I
************************* (    12)  12.1 PCT
I    0.30-0.40
I
I
********************************* (    16)  16.2 PCT
I    0.40-0.50
I
I
*************************************************** (    25)  25.3 PCT
I    0.50-0.60
I
I
********************************************* (    22)  22.2 PCT
I    0.60-0.70
I
I
***************** (     8)   8.1 PCT
I    0.70-0.80
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

            EXEMPLAR SENSITIVITY-ANDERSON

I
********************* (     7)   7.1 PCT
I    -1.00-0.00
I
I
*************** (     5)   5.1 PCT
I    0.00-0.10
I
I
*************************************** (    13)  13.1 PCT
I    0.10-0.20
I
I
********************************************* (    15)  15.2 PCT
I    0.20-0.30
I
I
****************************************************** (    18)  18.2 PCT
I    0.30-0.40
I
I
*************************************************************************** (    25)  25.3 PCT
I    0.40-0.50
I
I
************************************ (    12)  12.1 PCT
I    0.50-0.60
I
I
************ (     4)   4.0 PCT
I    0.60-0.70
I
I..............I..............I..............I..............I..............I..............I..............I..............I..............I..............I
0              5             10             15             20             25             30             35             40             45             50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

EXEMPLAR SENSITIVITY-COLE

```
*****   (     2)   2.0 PCT
I   0.20-0.30
I
*******  (     3)   3.0 PCT
I   0.30-0.40
I
***********************  (    11)   11.1 PCT
I   0.40-0.50
I
***********************  (    11)   11.1 PCT
I   0.50-0.60
I
************************************************  (    24)   24.2 PCT
I   0.60-0.70
I
************************************************************************  (    36)   36.4 PCT
I   0.70-0.80
I
*************************  (    12)   12.1 PCT
I   0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS H-PRE

EXEMPLAR SENSITIVITY-HALL

```
I
*** (     1)   1.0 PCT
I    -1.00-0.20
I
***** (     2)   1.9 PCT
I    0.20-0.30
I
******* (     3)   2.9 PCT
I    0.30-0.40
I
******* (     3)   2.9 PCT
I    0.40-0.50
I
******************************************** (    16)  15.2 PCT
I    0.50-0.60
I
********************************************** (    17)  16.2 PCT
I    0.60-0.70
I
********************************************************************************************************************************** (    49)  46.7 PCT
I    0.70-0.80
I
************************************ (    13)  12.4 PCT
I    0.80-0.90
I
*** (     1)   1.0 PCT
I    0.90-1.00
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS R-PPF

                    EXEMPLAR SENSITIVITY-MURDOCK

I
****** (     3)   2.9 PCT
I    -1.00-0.00
I
I
****** (     3)   2.9 PCT
I    0.00-0.10
I
I
**************** (     8)   7.6 PCT
I    0.10-0.20
I
I
******************************** (    16)   15.2 PCT
I    0.20-0.30
I
I
******************************************** (    22)   21.0 PCT
I    0.30-0.40
I
I
************************************************** (    25)   23.8 PCT
I    0.40-0.50
I
I
************************************ (    18)   17.1 PCT
I    0.50-0.60
I
I
************** (     7)   6.7 PCT
I    0.60-0.70
I
I
****** (     3)   2.9 PCT
I    0.70-0.80
I
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5         10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS H-PRF

            EXEMPLAR SENSITIVITY-ANDERSON

I
I****** (     3)   2.9 PCT
I    -1.00-0.00
I
I******** (     4)   3.8 PCT
I    0.00-0.10
I
I****************** (     9)   8.6 PCT
I    0.10-0.20
I
I**************************************** (    20)  19.0 PCT
I    0.20-0.30
I
I****************************************************** (    27)  25.7 PCT
I    0.30-0.40
I
I************************************************ (    24)  22.9 PCT
I    0.40-0.50
I
I**************************** (    14)  13.3 PCT
I    0.50-0.60
I
I****** (     3)   2.9 PCT
I    0.60-0.70
I
I** (     1)   1.0 PCT
I    0.70-0.80
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRE

        EXEMPLAR SENSITIVITY-COLE

I
** (     1)   1.0 PCT
I    -1.00-0.00
I
I
****** (     3)   2.9 PCT
I     0.20-0.30
I
I
****** (     3)   2.9 PCT
I     0.30-0.40
I
I
**************** (     8)   7.6 PCT
I     0.40-0.50
I
I
****************************************  (    20)  19.0 PCT
I     0.50-0.60
I
I
******************************************************************** (    34)  32.4 PCT
I     0.60-0.70
I
I
******************************************************** (    28)  26.7 PCT
I     0.70-0.80
I
I
**************** (     8)   7.6 PCT
I     0.80-0.90
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - APRF

EXEMPLAR SENSITIVITY

HALL

MEAN	6.343	STD ERROR	.120	MEDIAN	6.387
MODE	6.000	STD DEV	1.197	VARIANCE	1.432
KURTOSIS	-.155	SKEWNESS	-.255	RANGE	6.000
MINIMUM	3.000	MAXIMUM	9.000		

MURDOCK

MEAN	5.350	STD ERROR	.165	MEDIAN	5.726
MODE	6.000	STD DEV	1.842	VARIANCE	3.394
KURTOSIS	.116	SKEWNESS	-.784	RANGE	8.000
MINIMUM	0.000	MAXIMUM	8.000		

ANDERSON

MEAN	3.768	STD ERROR	.184	MEDIAN	4.028
MODE	5.000	STD DEV	1.829	VARIANCE	3.343
KURTOSIS	-.576	SKEWNESS	.408	RANGE	7.000
MINIMUM	0.000	MAXIMUM	7.000		

COLE

MEAN	7.101	STD ERROR	.142	MEDIAN	7.438
MODE	8.000	STD DEV	1.418	VARIANCE	2.010
KURTOSIS	.147	SKEWNESS	-.849	RANGE	6.000
MINIMUM	5.000	MAXIMUM	9.000		

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - BPRF

EXEMPLAR SENSITIVITY

HALL

MEAN	7.305	STD ERROR	.147	MEDIAN	7.714
MODE	8.000	STD DEV	1.507	VARIANCE	2.272
KURTOSIS	4.891	SKEWNESS	-1.781	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

MURDOCK

MEAN	4.390	STD ERROR	.170	MEDIAN	4.520
MODE	5.000	STD DEV	1.740	VARIANCE	3.029
KURTOSIS	-.147	SKEWNESS	-.327	RANGE	8.000
MINIMUM	0.000	MAXIMUM	8.000		

ANDERSON

MEAN	4.029	STD ERROR	.155	MEDIAN	4.111
MODE	4.000	STD DEV	1.590	VARIANCE	2.528
KURTOSIS	.064	SKEWNESS	-.321	RANGE	8.000
MINIMUM	0.000	MAXIMUM	8.000		

COLE

MEAN	6.810	STD ERROR	.146	MEDIAN	7.015
MODE	7.000	STD DEV	1.494	VARIANCE	2.233
KURTOSIS	3.434	SKEWNESS	-1.375	RANGE	9.000
MINIMUM	0.000	MAXIMUM	9.000		

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

GROUP CONST... SENSITIVITY ...

```
I
***  (    1)   1.0 PCT
     0.30-0.40
I
*****  (    2)   2.0 PCT
     0.40-0.50
I
*********  (    4)   4.0 PCT
     0.50-0.60
I
*****  (    2)   2.0 PCT
     0.60-0.70
I
*****************************************  (   19)  19.2 PCT
     0.70-0.80
I
*******************************************************************************************  (   45)  45.5 PCT
     0.80-0.90
I
*****************************************************  (   26)  26.3 PCT
     0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQ ENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

Interval	Frequency	Percent
-1.00-0.00	1	1.0 PCT
0.50-0.60	7	7.1 PCT
0.60-0.70	10	10.1 PCT
0.70-0.80	16	16.2 PCT
0.80-0.90	38	38.4 PCT
0.90-1.00	27	27.3 PCT

FREQUENCY 0 5 10 15 20 25 30 35 40 45 50

```
DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

            GROUP CONSENSUS SENSITIVITY - ANDERSON

*******  (    3)   3.0 PCT
I    -1.00-0.00
I
I
***  (    1)   1.0 PCT
I    0.00-0.10
I
I
*****  (    2)   2.0 PCT
I    0.10-0.20
I
I
***  (    1)   1.0 PCT
I    0.20-0.30
I
I
*******  (    3)   3.0 PCT
I    0.30-0.40
I
I
*******  (    3)   3.0 PCT
I    0.40-0.50
I
I
***  (    1)   1.0 PCT
I    0.50-0.60
I
I
***************************************  (   19)  19.2 PCT
I    0.60-0.70
I
I
*********************************************************  (   28)  28.3 PCT
I    0.70-0.80
I
I
*********************************************************************  (   34)  34.3 PCT
I    0.80-0.90
I
I
*********  (    4)   4.0 PCT
I    0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS A-PRF

GROUP CONSENSUS SENSITIVITY SCORE

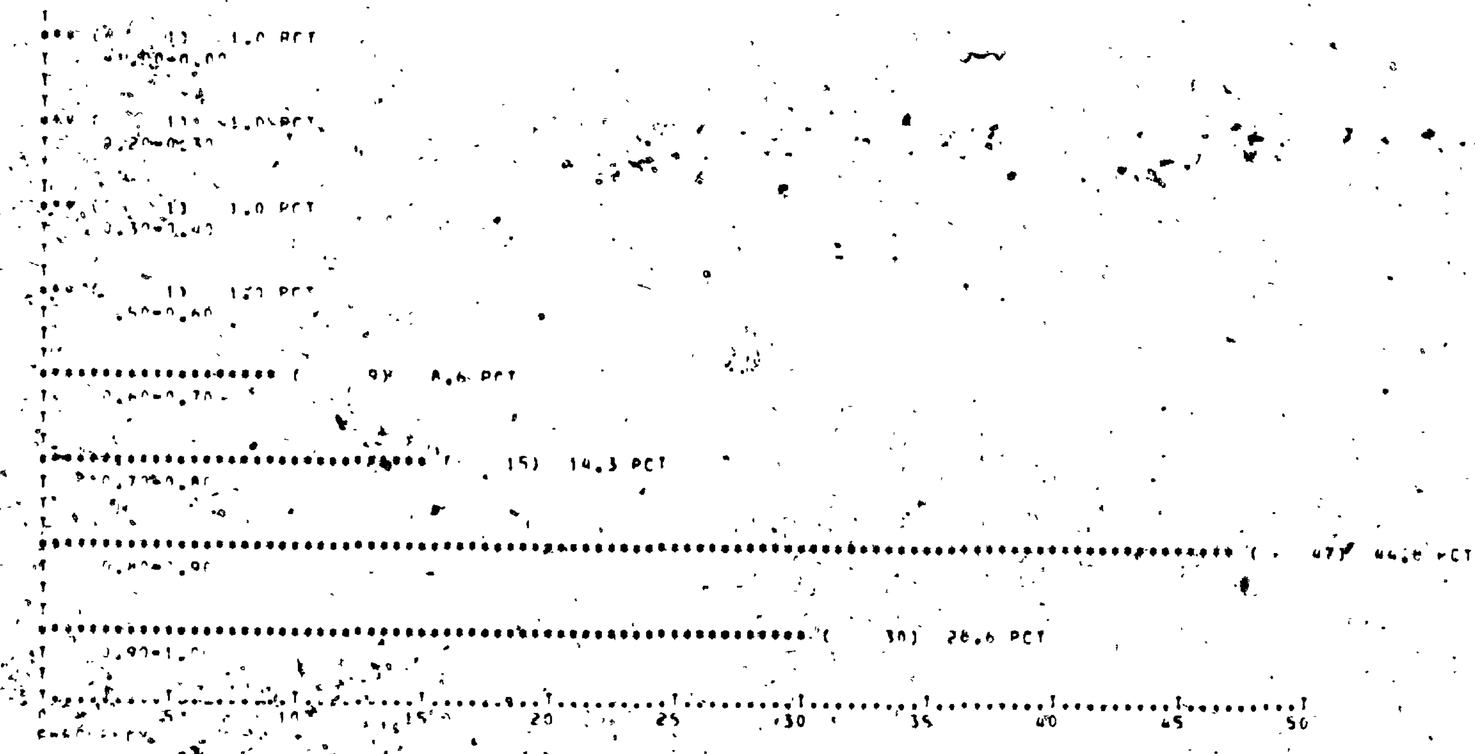
0.40-0.50	(1)	1.0 PCT
0.50-0.60	(2)	2.0 PCT
0.60-0.70	(4)	4.0 PCT
0.70-0.80	(18)	18.2 PCT
0.80-0.90	(36)	36.4 PCT
0.90-1.00	(38)	38.4 PCT

0 5 10 15 20 25 30 35 40 45 50

FREQUENCY

DISTRIBUTION OF SUBJECT PARAMETERS H-PRE

GROUP CONSENSUS SENSITIVITY - HALL



```
DISTRIBUTION OF SUBJECT PARAMETERS R=PRF

          GROUP CONSENSUS SENSITIVITY - MURDOCK

I
*** (    1)   1.0 PCT
I   -1.00-0.00
I
*** (    1)   1.0 PCT
I    0.20-0.30
I
*** (    1)   1.0 PCT
I    0.40-0.50
I
*********** (    5)   4.8 PCT
I    0.50-0.60
I
**************** (    8)   7.6 PCT
I    0.60-0.70
I
******************************** (   16)  15.2 PCT
I    0.70-0.80
I
********************************************************************************************** (   46)  43.8 PCT
I    0.80-0.90
I
******************************************************** (   27)  25.7 PCT
I    0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0         5        10        15        20        25        30        35        40        45        50
FREQUENCY
```

DISTRIBUTION OF SUBJECT PARAMETERS R=PHF

GROUP CONSENSUS SENSITIVITY - ANDERSON

Interval	Frequency	Percent
0.10-0.20	1	1.0 PCT
0.30-0.40	1	1.0 PCT
0.40-0.50	4	3.8 PCT
0.50-0.60	2	1.9 PCT
0.60-0.70	17	16.2 PCT
0.70-0.80	23	21.9 PCT
0.80-0.90	49	46.7 PCT
0.90-1.00	8	7.6 PCT

FREQUENCY (axis: 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50)

```
DISTRIBUTION OF SUBJECT PARAMETERS H=PRF

         GROUP CONTENT OF [illegible] - CELT

I
**          (    2)    1.9 PCT
I     0.20-0.30
I
I
*  (        1)    1.0 PCT
I     0.30-0.40
I
I
*  (        1)    1.0 PCT
I     0.40-0.50
I
I
****  (      4)    3.8 PCT
I     0.50-0.60
I
I
*****  (      5)    4.8 PCT
I     0.60-0.70
I
I
***********  (     11)    10.5 PCT
I     0.70-0.80
I
I
**********************************************************  (     58)    55.2 PCT
I     0.80-0.90
I
I
***********************  (     23)    21.9 PCT
I     0.90-1.00
I
I.........I.........I.........I.........I.........I.........I.........I.........I.........I.........I
0        10        20        30        40        50        60        70        80        90       100
FREQUENCY
```

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - APRF

GROUP CONSENSUS

HALL

MEAN	8.778	STD ERROR	.121	MEDIAN	8.978
MODE	[illegible]	STD DEV	1.200	VARIANCE	1.440
KURTOSIS	3.223	SKEWNESS	-1.631	RANGE	6.000
MINIMUM	4.000	MAXIMUM	10.000		

MURDOCK

MEAN	8.606	STD ERROR	.148	MEDIAN	8.908
MODE	9.000	STD DEV	1.470	VARIANCE	2.160
KURTOSIS	10.218	SKEWNESS	-2.362	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

ANDERSON

MEAN	7.515	STD ERROR	.218	MEDIAN	8.089
MODE	9.000	STD DEV	2.173	VARIANCE	4.722
KURTOSIS	3.773	SKEWNESS	-1.985	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

COLE

MEAN	9.020	STD ERROR	.105	MEDIAN	9.181
MODE	10.000	STD DEV	1.040	VARIANCE	1.081
KURTOSIS	1.839	SKEWNESS	-1.244	RANGE	5.000
MINIMUM	5.000	MAXIMUM	10.000		

FREQUENCY DISTRIBUTION OF SUBJECT PARAMETERS - BPRF

GROUP CONSENSUS

HALL

MEAN	8.752	STD ERROR	.1[illegible]	MEDIAN	9.021
MODE	9.000	STD DEV	1.473	VARIANCE	2.169
KURTOSIS	12.877	SKEWNESS	-2.979	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

MURDOCK

MEAN	8.629	STD ERROR	[illegible]	MEDIAN	8.946
MODE	9.000	STD DEV	1.514	VARIANCE	2.293
KURTOSIS	10.130	SKEWNESS	-2.576	RANGE	10.000
MINIMUM	0.000	MAXIMUM	10.000		

ANDERSON

MEAN	8.210	STD ERROR	.132	MEDIAN	8.592
MODE	9.000	STD DEV	1.349	VARIANCE	1.821
KURTOSIS	4.057	SKEWNESS	-1.683	RANGE	8.000
MINIMUM	2.000	MAXIMUM	10.000		

COLE

MEAN	8.705	STD ERROR	.133	MEDIAN	8.891
MODE	9.000	STD DEV	1.365	VARIANCE	1.864
KURTOSIS	5.720	SKEWNESS	-2.212	RANGE	7.000
MINIMUM	3.000	MAXIMUM	10.000		

Intercorrelations of Elevation Parameters

Study A

	Murdock	Anderson	Cole
Hall	-.225	.037	.198
Murdock		.172	-.130
Anderson			.091

Study B

	Murdock	Anderson	Cole
Hall	-.002	.279	[illegible]25
Murdock		-.079	-.022
Anderson			.098

Intercorrelations of Dispersion Parameters

Study A

	Murdock	Anderson	Cole
Hall	.557	.530	.603
Murdock		.464	.544
Anderson			.651

Study B

	Murdock	Anderson	Cole
Hall	.751	.626	.557
Murdock		.612	.625
Anderson			.592